Prediction of final data with use of preliminary and/or revised data

Roberto, S. Mariano

Tanizaki, Hisashi

(Citation)
Journal of Forecasting,14(4):351-380

(Issue Date)
2006-11

(Resource Type)
journal article

(Version)
Accepted Manuscript

PREDICTION OF FINAL DATA WITH USE OF PRELIMINARY AND/OR REVISED DATA*

Roberto S. Mariano
University of Pennsylvania
Philadelphia, PA 19104, USA

Hisashi Tanizaki
Kobe University
Kobe 657, JAPAN

First Draft: February, 1994
Revised: September, 1994

Abstract: In the case of U.S. national accounts, the data are revised for the first few years and every decade, which implies that we do not really have the final data. In this paper, we aim to predict the final data, using the preliminary data and/or the revised data. The following predictors are introduced and derived from a context of the nonlinear filtering or smoothing problem, which are: (i) prediction of the final data of time t given the preliminary data up to time $t-1$, (ii) prediction of the final data of time t given the preliminary data up to time t, (iii) prediction of the final data of time t given the preliminary data up to time T, (iv) prediction of the final data of time t given the revised data up to time $t-1$ and the preliminary data up to time $t-1$, and (v) prediction of the final data of time t given the revised data up to time $t-1$ and the preliminary data up to time t. It is shown that (v) is the best predictor but not too different from (iii). The prediction problem is illustrated using U.S. per capita consumption data.

Key Words: Revision Process, Prediction, Kalman Filter, Monte-Carlo Integration Filter, Final Data, Revised Data, Preliminary Data

1 INTRODUCTION

There is a great amount of literature in evaluating expectation. Mariano and Brown (1983,1989) and Brown and Mariano (1984,1989) suggested using Monte-Carlo stochastic simulations for the expectation of a nonlinear function. Moreover, Kitagawa (1987) evaluated the expectation, applying numerical integration to the context of a nonlinear filtering problem. Also, Geweke (1989) and Shao (1989) applied the importance sampling to evaluating integration in the Bayesian framework. Tanizaki (1991, 1993) and Tanizaki and Mariano (1994) proposed the Monte-Carlo integration filter using the importance sampling, which is applied in this paper to prediction of final data based on preliminary data and/or revised data.

Numerous papers deal with the data revision process. Conrad and Corrado (1979) applied the Kalman filter to improve upon published preliminary estimates of monthly retail sales, using an ARIMA model. Howrey (1978, 1984) used the preliminary data in econometric forecasting and obtained substantial improvements in forecast accuracy when the preliminary and revised data are used optimally (also, see Harvey (1989)). In this paper, we consider the revision process in nonlinear and/or nonnormal cases.

In the case of annual data on U.S. national accounts, the preliminary data at the present time are reported at the beginning of the next year. The revision process is performed over a few years and every decade, which is shown in Table 1, taking an example of the nominal consumption data (U.S. Personal Consumption Expenditures, billion dollars). All the data in Table 1 are taken from *Economic Report of the President* (ERP), published from 1948 to 1994. Each column indicates the year when ERP is published, while each row represents the data of the corresponding year. The superscript p denotes the preliminary data and the superscript r implies the data revised in the year corresponding to each column,

*The authors are grateful to two anonymous referees for valuable comments and suggestions.

Table 1: Revision Process of US Personal Consumption Expenditures

	1948	1949	1950	1951	1952	1953	1954	1955	1956	1957
1946	143.7^r	147.4^r	147.8^r	146.9^r	146.9	146.9	146.9	146.6^r	146.6	146.6
1947	164.5^p	164.8^r	166.9^r	165.6^r	165.6	165.6	165.6	165.0^r	165.0	165.0
1948	NA	176.8^p	178.8^r	177.4^r	177.9^r	177.9	177.9	177.6^r	177.6	177.6
1949	NA	NA	178.5^p	178.8^r	180.2^r	180.6^r	180.6	180.6	180.6	180.6
1950	NA	NA	NA	190.8^p	193.6^r	194.3^r	194.6^r	194.0^r	194.0	194.0
1951	NA	NA	NA	NA	204.4^p	208.0^r	208.1^r	208.3^r	208.3	208.3
1952	NA	NA	NA	NA	NA	216.0^p	218.1^r	218.4^r	218.3^r	218.3
1953	NA	NA	NA	NA	NA	NA	229.9^p	230.1^r	230.6^r	230.5^r
1954	NA	NA	NA	NA	NA	NA	NA	234.0^p	236.5^r	236.5
1955	NA	NA	NA	NA	NA	NA	NA	NA	252.4^p	254.0^r
1956	NA	NA	NA	NA	NA	NA	NA	NA	NA	265.8^p

Source: *Economic Report of the President* (1948 – 1957)

and NA indicates that the data is not available (i.e., the data has not been published yet). For instance, take the consumption data of 1947 (the corresponding row in Table 1). The preliminary consumption of 1947 was reported in 1948 (i.e., 164.5) and it was revised in 1949 for the first time (i.e., 164.8). In 1950 and 1951, the second and the third revised data were published, respectively (i.e., 166.9 and 165.4). Since it was not revised in 1952, the consumption of 1947 is given by 165.4. Moreover, the data was revised as 165.0 in 1955, 165.4 in 1959, 160.7 in 1966, 161.7 in 1976, 161.9 in 1986, and 162.3 in 1992.

Thus, each data series is revised every year for the first few years and thereafter it is revised less frequently. The consumption data of 1947 has been revised nine times until 1994 since the preliminary consumption was published in 1948. In such a sense, we cannot really know the true final data because the data is revised forever, while the preliminary data is reported only once. Therefore, it might be possible to consider that the final data is unobservable, which leads to estimation of the final data given the preliminary data, i.e, $E(x_t^f|X_{t-1}^p)$, $E(x_t^f|X_t^p)$ and $E(x_t^f|X_T^p)$. x_t^f represents the final data of time t and X_s^p indicates the preliminary data up to time s for $s = t-1, t, T$, i.e., $X_s^p = \{x_s^p, x_{s-1}^p, \cdots, x_1^p\}$. T denotes the sample size. Moreover, in the ERP published at period $t+1$, the preliminary data of time t and the L-th revised data of time $t-L$ for $L = 1, 2, \cdots$ (i.e., the most recently revised data) are available. That is, when the preliminary data of time t (i.e., x_t^p) is reported, the L-th revised data of time $t-L$ for $L = 1, 2, \cdots$ (i.e., ${}_tX_{t-1}^f$) are also available, where ${}_tX_{t-1}^f$ denotes ${}_tX_{t-1}^f = \{{}_tx_{t-1}^f, \cdots, {}_tx_{t-L}^f, \cdots, {}_tx_1^f\}$.[1] Therefore, it is also useful to examine estimating the final data of time t given the revised data up to time $t-1$ and the preliminary data up to time $t-1$, i.e., $E(x_t^f|{}_tX_{t-1}^f, X_{t-1}^p)$, and the final data of time t given the revised data up to time $t-1$ and the preliminary data up to time t, i.e., $E(x_t^f|{}_tX_{t-1}^f, X_t^p)$.

Summarizing the above, in this paper, we consider the following predictors.[2]

(i) prediction of the final data of time t given the preliminary data up to time $t-1$, i.e., $E(x_t^f|X_{t-1}^p)$,

[1] ${}_tX_s^f$ denotes the revised data series up to time s found in the ERP published at time $t+1$ when the preliminary data of time t is reported in the ERP published at time $t+1$, i.e., ${}_tX_s^f = \{{}_tx_s^f, {}_tx_{s-1}^f, \cdots, {}_tx_1^f\}$. Therefore, the consumption data up to time s in the ERP published at time $t+1$ includes both the revised data and the not revised data (recall that the national account data are revised over a few years and every decade). ${}_{t+1}x_{t-L}^f$ is possibly equivalent to ${}_tx_{t-L}^f$ for some L.
For instance, the preliminary data of 1990 is 3658.1 (i.e., $x_t^p = 3658.1$ for $t = 1990$) and the revised data series published in 1991 is given by $\{3450.1, 3238.2, 3009.4, 2797.4, \cdots\}$ (i.e., ${}_tX_{t-1}^f = \{{}_tx_{t-1}^f, {}_tx_{t-2}^f, \cdots\} = \{3450.1, 3238.2, 3009.4, 2797.4, \cdots\}$ for $t = 1990$). Similarly, the preliminary data of 1989 is 3470.3 (i.e., $x_t^p = 3470.3$ for $t = 1990$) and the revised data series published in 1989 is given by $\{3235.1, 3010.8, 2797.4, 2629.0, \cdots\}$ (i.e., ${}_tX_{t-1}^f = \{{}_tx_{t-1}^f, {}_tx_{t-2}^f, \cdots\} = \{3235.1, 3010.8, 2797.4, 2629.0, \cdots\}$ for $t = 1989$). See Table 1.

[2] Each information set (i.e., X_{t-1}^p for (i), X_t^p for (ii), X_T^p for (iii), $\{{}_tX_{t-1}^f, X_{t-1}^p\}$ for (iv) and $\{{}_tX_{t-1}^f, X_t^p\}$ for (v)) may include the other exogenous variables for all t in the equations (1) and (2). Prediction (iv) is the most familiar predictor, which is broadly used. When we forecast a variable by econometric model, usually we use the most recently revised (or reported) data for each past time period. Note that the revised data ${}_tX_{t-1}^f$ and the preliminary data X_t^p are observed but the final data is unobservable when the ERP is published at time $t+1$.

Table 1: Revision Process of US Personal Consumption Expenditures — Continued

	1958	1959	1960	1961	1962	1963	1964	1965	1966	1967
1946	146.6	147.1^r	147.1	147.1	147.1	147.1	147.1	147.1	143.4^r	143.4
1947	165.0	165.4^r	165.4	165.4	165.4	165.4	165.4	165.4	160.7^r	160.7
1948	177.6	178.3^r	178.3	178.3	178.3	178.3	178.3	178.3	173.6^r	173.6
1949	180.6	181.2^r	181.2	181.2	181.2	181.2	181.2	181.2	176.8^r	176.8
1950	194.0	195.0^r	195.0	195.0	195.0	195.0	195.0	195.0	191.0^r	191.0
1951	208.3	209.8^r	209.8	209.8	209.8	209.8	209.8	209.8	206.3^r	206.3
1952	218.3	219.8^r	219.8	219.8	219.8	219.8	219.8	219.8	216.7^r	216.7
1953	230.5	232.6^r	232.6	232.6	232.6	232.6	232.6	232.6	230.0^r	230.0
1954	236.6^r	238.0^r	238.0	238.0	238.0	238.0	238.0	238.0	236.5^r	236.5
1955	254.4^r	256.9^r	256.9	256.9	256.9	256.9	256.9	256.9	254.4^r	254.4
1956	267.2^r	269.4^r	269.9^r	269.9	269.9	269.9	269.9	269.9	266.7^r	266.7
1957	280.4^p	284.4^r	284.8^r	285.2^r	285.2	285.2	285.2	285.2	281.4^r	281.4
1958	NA	290.6^p	293.0^r	293.5^r	293.2^r	293.2	293.2	293.2	290.1^r	290.1
1959	NA	NA	311.4^p	313.8^r	314.0^r	313.5^r	313.5	313.5	311.2^r	311.2
1960	NA	NA	NA	328.2^p	328.9^r	328.5^r	328.2^r	328.2	325.2^r	325.2
1961	NA	NA	NA	NA	339.2^p	338.1^r	336.8^r	337.3^r	335.2^r	335.2
1962	NA	NA	NA	NA	NA	356.7^p	355.4^r	356.8^r	355.1^r	355.1
1963	NA	NA	NA	NA	NA	NA	373.2^p	375.0^r	373.8^r	375.0^r
1964	NA	NA	NA	NA	NA	NA	NA	399.2^p	398.9^r	401.4^r
1965	NA	NA	NA	NA	NA	NA	NA	NA	428.5^p	431.5^r
1966	NA	NA	NA	NA	NA	NA	NA	NA	NA	465.0^p

Source: *Economic Report of the President* (1958 – 1967)

(ii) prediction of the final data of time t given the preliminary data up to time t, i.e., $E(x_t^f|X_t^p)$,

(iii) prediction of the final data of time t given the preliminary data up to time T, i.e., $E(x_t^f|X_T^p)$,

(iv) prediction of the final data of time t given the revised data up to time $t-1$ and the preliminary data up to time $t-1$, i.e., $E(x_t^f|{}_tX_{t-1}^f, X_{t-1}^p)$,

(v) prediction of the final data of time t given the revised data up to time $t-1$ and the preliminary data up to time t, i.e., $E(x_t^f|{}_tX_{t-1}^f, X_t^p)$.

The predictors (i) – (iii) can be derived from a filtering or smoothing technique. (v) is a combination of (ii) and (iv). Thus, by using the nonlinear filtering or smoothing technique, we examine the prediction problem of final data based on preliminary data and/or revised data.

2 SETUP OF THE MODEL

In this section, first, the revision problem is discussed in a general formulation. In the next section we show the more concrete example, taking the U.S. consumption data on national accounts.

Let x_t^f and x_t^p be the final data and the preliminary data, respectively. The model for the problem is specified as follows.[3]

Measurement equation $$q_t(x_t^p, x_t^f; \gamma) = \epsilon_t, \tag{1}$$

[3] In the context of the errors-in-variables, we can interpret the system (1) and (2) in another way. Usually, all the published data have measurement errors. If we estimate an econometric model with such data, an appropriate result cannot be obtained. Therefore, we can consider directly including the errors into the model. There, we may take x_t^f as a vector of the true unobservable variables, and x_t^p as a vector of the the actually obtained data corresponding to x_t^f. Thus, another interpretation is that the errors-in-variables problem is taken into account, where the exactly same system as (1) and (2) can be utilized.

Table 1: Revision Process of US Personal Consumption Expenditures — Continued

	1968	1969	1970	1971	1972	1973	1974	1975	1976	1977
1946	143.4	143.4	143.4	143.4	143.4	143.4	143.4	143.4	143.8^r	143.8
1947	160.7	160.7	160.7	160.7	160.7	160.7	160.7	160.7	161.7^r	161.7
1948	173.6	173.6	173.6	173.6	173.6	173.6	173.6	173.6	174.7^r	174.7
1949	176.8	176.8	176.8	176.8	176.8	176.8	176.8	176.8	178.1^r	178.1
1950	191.0	191.0	191.0	191.0	191.0	191.0	191.0	191.0	192.0^r	192.0
1951	206.3	206.3	206.3	206.3	206.3	206.3	206.3	206.3	207.1^r	207.1
1952	216.7	216.7	216.7	216.7	216.7	216.7	216.7	216.7	217.1^r	217.1
1953	230.0	230.0	230.0	230.0	230.0	230.0	230.0	230.0	229.7^r	229.7
1954	236.5	236.5	236.5	236.5	236.5	236.5	236.5	236.5	235.8^r	235.8
1955	254.4	254.4	254.4	254.4	254.4	254.4	254.4	254.4	253.7^r	253.7
1956	266.7	266.7	266.7	266.7	266.7	266.7	266.7	266.7	266.0^r	266.0
1957	281.4	281.4	281.4	281.4	281.4	281.4	281.4	281.4	280.4^r	280.4
1958	290.1	290.1	290.1	290.1	290.1	290.1	290.1	290.1	289.5^r	289.5
1959	311.2	311.2	311.2	311.2	311.2	311.2	311.2	311.2	310.8^r	310.8
1960	325.2	325.2	325.2	325.2	325.2	325.2	325.2	325.2	324.9^r	324.9
1961	335.2	335.2	335.2	335.2	335.2	335.2	335.2	335.2	335.0^r	335.0
1962	355.1	355.1	355.1	355.1	355.1	355.1	355.1	355.1	355.2^r	355.2
1963	375.0	375.0	375.0	375.0	375.0	375.0	375.0	375.0	374.6^r	374.6
1964	401.2^r	401.2	401.2	401.2	401.2	401.2	401.2	401.2	400.4^r	400.4
1965	433.1^r	432.8^r	432.8	432.8	432.8	432.8	432.8	432.8	430.2^r	430.2
1966	465.9^r	465.5^r	466.3^r	466.3	466.3	466.3	466.3	466.3	464.8^r	464.8
1967	491.6^p	492.2^r	492.3^r	492.1^r	492.1	492.1	492.1	492.1	490.4^r	490.4
1968	NA	533.7^p	536.6^r	535.8^r	536.2^r	536.2	536.2	536.2	535.9^r	535.9
1969	NA	NA	576.0^p	577.5^r	579.6^r	579.5^r	579.5	579.5	579.7^r	579.7
1970	NA	NA	NA	616.8^p	615.8^r	616.8^r	617.6^r	617.6	618.8^r	618.8
1971	NA	NA	NA	NA	662.2^p	664.9^r	667.2^r	667.1^r	668.2^r	668.2
1972	NA	NA	NA	NA	NA	721.1^p	726.5^r	729.0^r	733.0^r	733.0
1973	NA	NA	NA	NA	NA	NA	805.0^p	805.2^r	808.5^r	809.9^r
1974	NA	NA	NA	NA	NA	NA	NA	877.0^p	885.9^r	887.5^r
1975	NA	NA	NA	NA	NA	NA	NA	NA	963.2^p	973.2^r
1976	NA	NA	NA	NA	NA	NA	NA	NA	NA	1078.6^p

Source: *Economic Report of the President* (1968 – 1977)

Table 1: Revision Process of US Personal Consumption Expenditures — Continued

	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987
1946	143.8	143.8	143.8	143.8	143.8	143.8	143.8	143.8	143.9^r	143.9
1947	161.7	161.7	161.7	161.7	161.7	161.7	161.7	161.7	161.9^r	161.9
1948	174.7	174.7	174.7	174.7	174.7	174.7	174.7	174.7	174.9^r	174.9
1949	178.1	178.1	178.1	178.1	178.1	178.1	178.1	178.1	178.3^r	178.3
1950	192.0	192.0	192.0	192.0	192.0	192.0	192.0	192.0	192.1^r	192.1
1951	207.1	207.1	207.1	207.1	207.1	207.1	207.1	207.1	208.1^r	208.1
1952	217.1	217.1	217.1	217.1	217.1	217.1	217.1	217.1	219.1^r	219.1
1953	229.7	229.7	229.7	229.7	229.7	229.7	229.7	229.7	232.6^r	232.6
1954	235.8	235.8	235.8	235.8	235.8	235.8	235.8	235.8	239.8^r	239.8
1955	253.7	253.7	253.7	253.7	253.7	253.7	253.7	253.7	257.9^r	257.9
1956	266.0	266.0	266.0	266.0	266.0	266.0	266.0	266.0	270.6^r	270.6
1957	280.4	280.4	280.4	280.4	280.4	280.4	280.4	280.4	285.3^r	285.3
1958	289.5	289.5	289.5	289.5	289.5	289.5	289.5	289.5	294.6^r	294.6
1959	310.8	310.8	310.8	310.8	310.8	310.8	310.8	310.8	316.3^r	316.3
1960	324.9	324.9	324.9	324.9	324.9	324.9	324.9	324.9	330.7^r	330.7
1961	335.0	335.0	335.0	335.0	335.0	335.0	335.0	335.0	341.1^r	341.1
1962	355.2	355.2	355.2	355.2	355.2	355.2	355.2	355.2	361.9^r	361.9
1963	374.6	374.6	374.6	374.6	374.6	374.6	374.6	374.6	381.7^r	381.7
1964	400.4	400.4	400.4	400.5^r	400.5	400.5	400.5	400.5	409.3^r	409.3
1965	430.2	430.2	430.2	430.4^r	430.4	430.4	430.4	430.4	440.7^r	440.7
1966	464.8	464.8	464.8	465.1^r	465.1	465.1	465.1	465.1	477.3^r	477.3
1967	490.4	490.4	490.4	490.3^r	490.3	490.3	490.3	490.3	503.6^r	503.6
1968	535.9	535.9	535.9	536.9^r	536.9	536.9	536.9	536.9	552.5^r	552.5
1969	579.7	579.7	579.7	581.8^r	581.8	581.8	581.8	581.8	597.9^r	597.9
1970	618.8	618.8	618.8	621.7^r	621.7	621.7	621.7	621.7	640.0^r	640.0
1971	668.2	668.2	668.2	672.2^r	672.2	672.2	672.2	672.2	691.6^r	691.6
1972	733.0	733.0	733.0	737.1^r	737.1	737.1	737.1	737.1	757.6^r	757.6
1973	809.9	809.9	809.9	812.0^r	812.0	812.0	812.0	812.0	837.2^r	837.2
1974	889.6^r	889.6	889.6	888.1^r	888.1	888.1	888.1	888.1	916.5^r	916.5
1975	980.4^r	979.1^r	979.1	976.4^r	976.4	976.4	976.4	976.4	1012.8^r	1012.8
1976	1094.0^r	1090.2^r	1089.9^r	1084.3^r	1084.3	1084.3	1084.3	1084.3	1129.3^r	1129.3
1977	1210.1^p	1206.5^r	1210.0^r	1205.5^r	1205.5	1204.4^r	1204.4	1204.4	1257.2^r	1257.2
1978	NA	1339.7^p	1350.8^r	1348.7^r	1348.7	1346.5^r	1346.5	1346.5	1403.5^r	1403.5
1979	NA	NA	1509.8^p	1510.9^r	1510.9	1507.2^r	1507.2	1507.2	1566.8^r	1566.8
1980	NA	NA	NA	1670.1^p	1672.8^r	1667.2^r	1668.1^r	1668.1	1732.6^r	1732.6
1981	NA	NA	NA	NA	1858.1^p	1843.2^r	1857.2^r	1849.1^r	1915.1^r	1915.1
1982	NA	NA	NA	NA	NA	1972.0^p	1991.9^r	1984.9^r	2050.7^r	2050.7
1983	NA	NA	NA	NA	NA	NA	2158.6^p	2155.9^r	2229.3^r	2234.5^r
1984	NA	NA	NA	NA	NA	NA	NA	2342.3^p	2423.0^r	2428.2^r
1985	NA	NA	NA	NA	NA	NA	NA	NA	2581.9^p	2600.5^r
1986	NA	NA	NA	NA	NA	NA	NA	NA	NA	2762.4^p

Source: *Economic Report of the President* (1978 – 1987)

Table 1: Revision Process of US Personal Consumption Expenditures — Continued

	1988	1989	1990	1991	1992	1993	1994	L_t	r_t	P_t	Y_t
1946	143.9	143.9	143.9	143.9	144.3^r	144.3	144.3	.141389	2.53	.185	159.3
1947	161.9	161.9	161.9	161.9	162.3^r	162.3	162.3	.144126	2.61	.205	169.1
1948	174.9	174.9	174.9	174.9	175.4^r	175.4	175.4	.146631	2.82	.216	188.4
1949	178.3	178.3	178.3	178.3	178.9^r	178.9	178.9	.149188	2.66	.215	188.1
1950	192.1	192.1	192.1	192.1	192.7^r	192.7	192.7	.151684	2.62	.220	207.7
1951	208.1	208.1	208.1	208.1	208.7^r	208.7	208.7	.154287	2.86	.233	228.1
1952	219.1	219.1	219.1	219.1	219.7^r	219.7	219.7	.156954	2.96	.238	240.2
1953	232.6	232.6	232.6	232.6	233.5^r	233.5	233.5	.159565	3.20	.243	255.5
1954	239.8	239.8	239.8	239.8	240.7^r	240.7	240.7	.162391	2.90	.244	261.2
1955	257.9	257.9	257.9	257.9	259.1^r	259.1	259.1	.165275	3.06	.247	279.9
1956	270.6	270.6	270.6	270.6	271.9^r	271.9	271.9	.168221	3.36	.252	298.8
1957	285.3	285.3	285.3	285.3	286.7^r	286.7	286.7	.171274	3.89	.260	315.2
1958	294.6	294.6	294.6	294.6	296.3^r	296.3	296.3	.174141	3.79	.264	326.3
1959	316.3	316.3	316.3	316.3	318.1^r	318.1	318.1	.177073	4.38	.270	346.7
1960	330.7	330.7	330.7	330.7	332.4^r	332.4	332.4	.180760	4.41	.275	360.5
1961	341.1	341.1	341.1	341.1	343.5^r	343.5	343.5	.183742	4.35	.277	376.2
1962	361.9	361.9	361.9	361.9	364.4^r	364.4	364.4	.186590	4.33	.282	398.7
1963	381.7	381.7	381.7	381.7	384.2^r	384.2	384.2	.189300	4.26	.286	418.4
1964	409.3	409.3	409.3	409.3	412.5^r	412.5	412.5	.191927	4.40	.291	454.7
1965	440.7	440.7	440.7	440.7	444.6^r	444.6	444.6	.194347	4.49	.297	491.0
1966	477.3	477.3	477.3	477.3	481.6^r	481.6	481.6	.196599	5.13	.306	530.7
1967	503.6	503.6	503.6	503.6	509.3^r	509.3	509.3	.198752	5.51	.314	568.6
1968	552.5	552.5	552.5	552.5	559.1^r	559.1	559.1	.200745	6.18	.327	617.8
1969	597.9	597.9	597.9	597.9	603.7^r	603.7	603.7	.202736	7.03	.341	663.8
1970	640.0	640.0	640.0	640.0	646.5^r	646.5	646.5	.205089	8.04	.356	722.0
1971	691.6	691.6	691.6	691.6	700.3^r	700.3	700.3	.207692	7.39	.374	784.9
1972	757.6	757.6	757.6	757.6	767.8^r	767.8	767.8	.209924	7.21	.388	848.5
1973	837.2	837.2	837.2	837.2	848.1^r	848.1	848.1	.211939	7.44	.410	958.1
1974	916.5	916.5	916.5	916.5	927.7^r	927.7	927.7	.213898	8.57	.452	1046.5
1975	1012.8	1012.8	1012.8	1012.8	1024.9^r	1024.9	1024.9	.215981	8.83	.489	1150.9
1976	1129.3	1129.3	1129.3	1129.3	1143.1^r	1143.1	1143.1	.218086	8.43	.518	1264.0
1977	1257.2	1257.2	1257.2	1257.2	1271.5^r	1271.5	1271.5	.220289	8.02	.554	1391.3
1978	1403.5	1403.5	1403.5	1403.5	1421.2^r	1421.2	1421.2	.222629	8.73	.594	1567.8
1979	1566.8	1566.8	1566.8	1566.8	1583.7^r	1583.7	1583.7	.225106	9.63	.647	1753.0
1980	1732.6	1732.6	1732.6	1732.6	1748.1^r	1748.1	1748.1	.227715	11.94	.714	1952.9
1981	1915.1	1915.1	1915.1	1915.1	1926.2^r	1926.2	1926.2	.229989	14.17	.778	2174.5
1982	2050.7	2050.7	2050.7	2050.7	2059.2^r	2059.2	2059.2	.232201	13.79	.822	2319.6
1983	2234.5	2234.5	2234.5	2234.5	2257.5^r	2257.5	2257.5	.234326	12.04	.862	2493.7
1984	2430.5^r	2430.5	2430.5	2430.5	2460.3^r	2460.3	2460.3	.236393	12.71	.896	2759.5
1985	2629.4^r	2629.0^r	2629.0	2629.0	2667.4^r	2667.4	2667.4	.238510	11.37	.931	2943.0
1986	2799.8^r	2807.5^r	2797.4^r	2797.4	2850.6^r	2850.6	2850.6	.240691	9.02	.960	3131.5
1987	2966.0^p	3012.1^r	3010.8^r	3009.4^r	3052.2^r	3052.2	3052.2	.242860	9.38	1.000	3289.5
1988	NA	3226.0^p	3235.1^r	3238.2^r	3296.1^r	3296.1	3296.1	.245093	9.71	1.042	3548.2
1989	NA	NA	3470.3^p	3450.1^r	3517.9^r	3523.1^r	3523.1	.247397	9.26	1.093	3787.0
1990	NA	NA	NA	3658.1^p	3742.6^r	3748.4^r	3761.2^r	.249951	9.32	1.149	4050.5
1991	NA	NA	NA	NA	3886.8^p	3887.7^r	3906.4^r	.252699	8.77	1.199	4230.5
1992	NA	NA	NA	NA	NA	4093.9^p	4139.9^r	.255472	8.14	1.239	4500.2
1993	NA	NA	NA	NA	NA	NA	4390.6^p	.258256	7.22	1.272	4706.0

Source: *Economic Report of the President* (1988 – 1994)

Transition equation $\qquad f_t(x_t^f, x_{t-1}^f; \theta) = \eta_t, \qquad (2)$

$$\begin{pmatrix} \epsilon_t \\ \eta_t \end{pmatrix} \sim \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \Sigma_\epsilon & 0 \\ 0 & \Sigma_\eta \end{pmatrix} \right),$$

for $t = 1, 2, \cdots, T$, where γ and θ are the unknown parameters to be estimated. The functional forms of q_t and f_t are assumed to be known.[4] ϵ_t and η_t are the error terms, which are assumed to be mutually independently and identically distributed, and not to be necessarily multivariate-normal.

The preliminary data, x_t^p, is closely related to the final data, x_t^f, as the measurement equation (1), because both are originally the same data (see Conrad and Corrado (1979) and Howrey (1978, 1984)).

Equation (2) represents an economic relationship, whish is related to the final data, because the preliminary data has the errors (i.e., ϵ_t) different from the errors (i.e., η_t) derived from an economic model. It might be appropriate that the unobservable final data x_t^f, rather than the observed preliminary data x_t^p, is related to the economic relationship.

The nonlinear filtering techniques are performed in order to obtain the prediction of x_t^f. The predictors can be derived as follows.

(i) $E(x_t^f|X_{t-1}^p)$ and (ii) $E(x_t^f|X_t^p)$:

Each data series is revised over a few years and every years, as shown in Table 1. For the predictors (i) and (ii), therefore, we consider that we do not have the true final data because they are revised forever. The preliminary data are reported only once. A natural idea is estimating the final data x_t^f given the preliminary data up to time t or $t-1$, which leads to a filtering problem.

The information set, i.e., X_{t-1}^p for (i) and X_t^p for (ii), is observable, but x_t^f is not observed. Let $P(\cdot|\cdot)$ be a conditional density function. The distribution-based filtering algorithm corresponding to the predictors (i) and (ii) is given by the following recursions (see, for example, Kitagawa (1987), Kramer and Sorenson (1988), Harvey (1989) and Tanizaki (1991,1993)):

$$P(x_t^f|X_{t-1}^p) = \begin{cases} P(x_t^f|x_{t-1}^f), & \text{if } t=1, \\ \displaystyle\int P(x_t^f|x_{t-1}^f) P(x_{t-1}^f|X_{t-1}^p) \mathrm{d}x_{t-1}^f, & \text{otherwise}, \end{cases} \tag{3}$$

$$P(x_t^f|X_t^p) = \frac{P(x_t^p|x_t^f) P(x_t^f|X_{t-1}^p)}{\displaystyle\int P(x_t^p|x_t^f) P(x_t^f|X_{t-1}^p) \mathrm{d}x_t^f}, \tag{4}$$

The initial density $P(x_1^f|X_0^p)$ is given by $P(x_1^f|x_0^f)$ if x_0^f is assumed to be known.[5] $P(x_t^f|x_{t-1}^f)$ in equation (3) is computed from the transition equation (2) if the distribution function of η_t is specified, and similarly $P(x_t^p|x_t^f)$ in equation (4) is derived from the measurement equation (1) given a specific distribution for ϵ_t. Thus, $P(x_t^f|X_{t-1}^p)$ and $P(x_t^f|X_t^p)$ are computed recursively, based on $P(x_t^f|x_{t-1}^f)$ and $P(x_t^p|x_t^f)$.[6]

In the case where the unknown parameters γ and θ are included in the system, the following likelihood function is maximized with respect to γ and θ:

$$L(\gamma, \theta; X_t^p) \equiv \prod_{t=1}^{T} P(x_t^p|X_{t-1}^p), \tag{5}$$

[4]The vector functions q_t and f_t may depend on the other exogenous variables, which are omitted from the two equations (1) and (2).

[5]If the initial value x_0^f is assumed to be stochastic, the case $t=1$ of equation (3) is written as

$$P(x_1^f|X_0^p) = \int P(x_1^f|x_0^f) P(x_0^f) \mathrm{d}x_0^f,$$

where $P(x_0^f)$ is a distribution function of the initial value x_0^f.

[6]Equations (1) and (2) may depend on the past information set X_{t-1}^p. In this case, $P(x_t^f|x_{t-1}^f)$ and $P(x_t^p|x_t^f)$ are replaced by $P(x_t^f|x_{t-1}^f, X_{t-1}^p)$ and $P(x_t^p|x_t^f, X_{t-1}^p)$.

for $t = 1, 2, \cdots, T$, where $P(x_t^p|X_{t-1}^p) = \int P(x_t^p|x_t^f)P(x_t^f|X_{t-1}^p)\mathrm{d}x_t^f$, which is the denominator of equation (4).

(iii) $E(x_t^f|X_T^p)$:

We can consider estimating the final data of time t given information set available at time T, i.e., $E(x_t^f|X_T^p)$ for $t = 1, \cdots, T$. Thus, the predictor (iii) estimates the past final data given preliminary data available at current period, which is equivalent to a smoothing problem. The distribution-based smoothing algorithm is represented by the following backward recursion (see, for example, Kitagawa (1987), Harvey (1989) and Tanizaki (1991,1993)):

$$P(x_{t-1}^f|X_T^p) = P(x_{t-1}^f|X_{t-1}^p)\int \frac{P(x_t^f|X_T^p)P(x_t^f|x_{t-1}^f)}{P(x_t^f|X_{t-1}^p)}\mathrm{d}x_t^f, \tag{6}$$

for $t = T, T-1, \cdots, 1$. $P(x_{t-1}^f|X_{t-1}^p)$ and $P(x_t^f|X_{t-1}^p)$ are obtained from equations (3) and (4).

(iv) $E(x_t^f|{}_tX_{t-1}^f, X_{t-1}^p)$ and (v) $E(x_t|{}_tX_{t-1}^f, X_t^p)$:

For the predictors (iv) and (v), the transition equation (2) in the state-space model is modified as follows.

$$\text{Transition equation} \qquad f_t(x_t^f, {}_tx_{t-1}^f; \theta) = \eta_t, \tag{2$'$}$$

for $t = 1, 2, \cdots, T$. The predictors (iv) and (v) are derived from the system (1) and (2′).

The information set for (iv) denotes $\{{}_tX_{t-1}^f, X_{t-1}^p\}$ which is the revised data series up to time $t-1$ and the preliminary data up to time $t-1$. For (v), the information set is given by $\{{}_tX_{t-1}^f, X_t^p\}$, which implies that the most recently revised data series up to time $t-1$ are utilized together with the preliminary data up to time t.

Thus, $E(x_t^f|{}_tX_{t-1}^f, X_{t-1}^p)$ predicts the final data of time t given the revised data up to $t-1$ and the preliminary data up to $t-1$, while $E(x_t^f|{}_tX_{t-1}^f, X_t^p)$ estimates the final data of time t given the revised data up to $t-1$ and the preliminary data up to t. In addition to the information set used in (i) and (ii), the most recently revised data are utilized for prediction. Therefore, it might be expected that (iv) and (v) show better predictors than (i) and (ii).

For the predictors (iv) and (v) based on equations (1) and (2′), the distributions $P(x_t^f|{}_tX_{t-1}^f, X_{t-1}^p)$ and $P(x_t^f|{}_tX_{t-1}^f, X_t^p)$ can be obtained as follows:

$$P(x_t^f|{}_tX_{t-1}^f, X_{t-1}^p) = P(x_t^f|{}_tx_{t-1}^f), \tag{7}$$

$$P(x_t^f|{}_tX_{t-1}^f, X_t^p) = \frac{P(x_t^p|x_t^f)P(x_t^f|{}_tX_{t-1}^f, X_{t-1}^p)}{\int P(x_t^p|x_t^f)P(x_t^f|{}_tX_{t-1}^f, X_{t-1}^p)\mathrm{d}x_t^f}, \tag{8}$$

for $t = 1, 2, \cdots, T$. The density formula (7) and (8) is given by replacing $P(x_t^f|X_{t-1}^p)$ and $P(x_t^f|X_t^p)$ by $P(x_t^f|{}_tX_{t-1}^f, X_{t-1}^p)$ [7] and $P(x_t^f|{}_tX_{t-1}^f, X_t^p)$ in the case $t = 1$ of equations (3) and (4). In the prediction (v), both the preliminary data up to time t and the revised data up to time $t-1$ are used for forecasting the final data of time t.

The information set for (v) is larger than that for (i), (ii) or (iv), because each information set is: X_{t-1}^p for (i), X_t^p for (ii), $\{{}_tX_{t-1}^f, X_{t-1}^p\}$ for (iv) and $\{{}_tX_{t-1}^f, X_t^p\}$ for (v). Therefore, we might expect that (v) is a better predictor of the final data than (i), (ii) and (iv).[8]

[7]Note that, since equation (2′) does not depend on X_{t-1}^p, $P(x_t^f|{}_tX_{t-1}^f, X_{t-1}^p)$ in equation (7) is equivalent to $P(x_t^f|{}_tX_{t-1}^f)$. Moreover, $P(x_t^f|{}_tX_{t-1}^f)$ reduces to $P(x_t^f|{}_tx_{t-1}^f)$ because transition equation (2′) depends on ${}_tx_{t-1}^f$ only, not $\{{}_tx_{t-2}^f, {}_tx_{t-3}^f, \cdots, {}_tx_1^f\}$. Accordingly, $P(x_t^f|{}_tX_{t-1}^f, X_{t-1}^p)$ is given by $P(x_t^f|{}_tx_{t-1}^f)$.

[8]Let us define each information set as $\Omega_1 = X_{t-1}^p$ for (i), $\Omega_2 = X_t^p$ for (ii), $\Omega_3 = X_T^p$ for (iii), $\Omega_4 = \{{}_tX_{t-1}^f, X_{t-1}^p\}$ for (iv) and $\Omega_5 = \{{}_tX_{t-1}^f, X_t^p\}$ for (v). Then, we have the following inclusion relations:

$$\Omega_1 \subset \Omega_2 \subset \Omega_3, \qquad \Omega_1 \subset \Omega_4 \subset \Omega_5, \qquad \Omega_2 \subset \Omega_5.$$

Therefore, it might be expected that (iii) or (v) is the best predictor.

Now, all the predictors were introduced in the same fashion. In the next section, we derive an optimal predictor in the sense of minimum mean square error for the following two cases: Section 3.1 Linear and Normal Case, and Section 3.2 Nonlinear and/or Nonnormal Case.

3 EVALUATION OF EACH EXPECTATION

In Section 2, we have introduced the predictors (i) – (v). In the case where the system is linear and normal, each optimal predictor in the sense of minimum mean square error can be rewritten as a linear algorithm. When the system (1) and (2) is nonlinear, however, there are two approaches for obtaining the estimates of the final data x_t^f. One is linearizing the nonlinear functions q_t and f_t by Taylor series expansions and applying the linearized functions to the conventional Kalman filter algorithm (see, for example, Wishner, Tabaczynski and Athans (1969) and Gelb (1974)). The most heuristic and easiest nonlinear filter is the extended Kalman filter, where the nonlinear functions are linearized by the first-order Taylor series expansion. However, it is well known that nonlinear filters based on Taylor series expansions give us biased estimators. Moreover, as Meinhold and Singpurwalla (1989) pointed out, the Kalman filter models under the normality assumption are not robust, and the posterior density becomes unrealistic when there is large difference between the observed data and the prior density. Therefore, the underlying density functions $P(x_t^f|X_t^p)$ and $P(x_t^f|X_{t-1}^p)$, rather than the nonlinear equations, should be approximated through the Gaussian sum (see Sorenson and Alspach (1971), Alspach and Sorenson (1972), and Anderson and Moore (1979)), numerical integration (see Kitagawa (1987) and Kramer and Sorenson (1988)), or Monte-Carlo integration (see Tanizaki (1991,1993) and Tanizaki and Mariano (1994)). Also, Carlin, Polson and Stoffer (1992) suggested applying an adaptive Monte-Carlo integration technique known as the Gibbs sampler to the density approximation. Needless to say, the density-based nonlinear filters do not require the normality and linearity assumptions, because the distributions of the state vector are exactly derived from the densities of ϵ_t and η_t and the functional forms of nonlinear measurement and transition equations (1) and (2). Thus, the density-based filters can be applied to the case where both the system is nonlinear and the error terms are non-Gaussian, and accordingly the obtained solutions are optimal.

3.1 Linear and Normal Case

First, consider the case where equations (1) and (2) are linear and error terms ϵ_t and η_t are normally distributed. The density algorithm represented by (3) and (4) reduces to the standard Kalman filter algorithm, which is derived from the first and second moments of the normal distributions.

In the linear and normal case, equations (1) and (2) are written as:

$$\begin{aligned} &\text{Measurement equation} \qquad && x_t^p = Z_t\, x_t^f + S_t\, \epsilon_t, \\ &\text{Transition equation} && x_t^f = T_t\, x_{t-1}^f + R_t\, \eta_t, \\ & && \begin{pmatrix} \epsilon_t \\ \eta_t \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \Sigma_\epsilon & 0 \\ 0 & \Sigma_\eta \end{pmatrix} \right), \end{aligned} \tag{9}$$

where Z_t, S_t, T_t and R_t are exogenous.[9] Under these assumptions, each expectation of (i) – (v) is represented as follows.

(i) $E(x_t^f|X_{t-1}^p)$ and (ii) $E(x_t^f|X_t^p)$:

Let us define: $E(x_t^f|X_s^p) \equiv x_{t|s}^f$ and $Var(x_t^f|X_s^p) \equiv \Sigma_{t|s}$ for $s = t-1, t$. It is well known that equations (3) and (4) reduces to the following standard Kalman filter algorithm when the system is linear and normal.

$$x_{t|t-1}^f = T_t x_{t-1|t-1}^f, \tag{10}$$

$$\Sigma_{t|t-1} = T_t \Sigma_{t-1|t-1} T_t' + R_t \Sigma_\eta R_t', \tag{11}$$

$$k_t = \Sigma_{t|t-1} Z_t' (Z_t \Sigma_{t|t-1} Z_t' + S_t \Sigma_\epsilon S_t')^{-1}, \tag{12}$$

$$x_{t|t}^f = x_{t|t-1}^f + k_t (x_t^p - Z_t x_{t|t-1}^f), \tag{13}$$

$$\Sigma_{t|t} = \Sigma_{t|t-1} - k_t (Z_t \Sigma_{t|t-1} Z_t' + S_t \Sigma_\epsilon S_t') k_t', \tag{14}$$

[9] Under normality assumption, Z_t, S_t, T_t and R_t may depend on the past information, i.e., X_{t-1}^p.

where the initial value is assumed to be known, i.e., $x^f_{0|0} = x^f_0$ and $\Sigma_{0|0} = 0$.[10] Once the initial values are given, $x^f_{t|s}$ and $\Sigma_{t|s}$ for $s = t-1, t$ are obtained recursively using the above algorithm (10) – (14). See Anderson and Moore (1979) and Harvey (1989). It is known that $x^f_{t|t}$ is optimal in the sense that it minimizes the mean square error. When the normality assumption is dropped, there is no longer any guarantee that the Kalman filter gives the conditional mean of the state vector. However, it is still an optimal estimator in the sense that it minimizes the mean square error within the class of all linear estimators, which implies that the Kalman filter under the normality assumption is the minimum mean square estimator and that the Kalman filter without the normality assumption is known as the minimum mean square linear estimator (see Harvey (1989)).

When Z_t, S_t, T_t and R_t depend on unknown parameters γ and θ, the log-likelihood function to be maximized with respect to the unknown parameters is given by:

$$\begin{aligned}\log L(\gamma,\theta;X^p_t) \quad = \quad & -\frac{T}{2}\log(2\pi) - \frac{1}{2}\sum_{t=1}^{T}\log|Z_t\Sigma_{t|t-1}Z_t' + S_t\Sigma_\epsilon S_t'| \\ & -\frac{1}{2}\sum_{t=1}^{T}(x^p_t - Z_t x^f_{t|t-1})'(Z_t\Sigma_{t|t-1}Z_t' + S_t\Sigma_\epsilon S_t')^{-1}(x^p_t - Z_t x^f_{t|t-1}), \qquad (15)\end{aligned}$$

which is derived from equation (5).

(iii) $E(x^f_t|X^p_T)$:

Let us define $x^f_{t|T} \equiv E(x^f_t|X^p_T)$ and $\Sigma_{t|T} \equiv Var(x^f_t|X^p_T)$. Under normality and linearity assumptions, the smoothing algorithm given by equation (6) reduces to the following backward recursion.

$$C_t = \Sigma_{t-1|t-1}T_t'\Sigma^{-1}_{t|t-1}, \qquad (16)$$

$$x^f_{t-1|T} = x^f_{t-1|t-1} + C_t(x_{t|T} - x_{t|t-1}), \qquad (17)$$

$$\Sigma_{t-1|T} = \Sigma_{t-1|t-1} + C_t(\Sigma_{t|T} - \Sigma_{t|t-1})C_t', \qquad (18)$$

where $t = T, T-1, \cdots, 1$.

(iv) $E(x^f_t|{}_tX^f_{t-1}, X^p_{t-1})$ and (v) $E(x^f_t|{}_tX^f_{t-1}, X^p_t)$:

For (iv) and (v), the transition equation (9) is modified as

Transition equation $\qquad x^f_t = T_t\ {}_tx^f_{t-1} + R_t\,\eta_t,$

where x^f_{t-1} in (9) is replaced by ${}_tx^f_{t-1}$.

Let us re-define $x^f_{t|s} \equiv E(x^f_t|{}_tX^f_{t-1}, X^p_s)$ and $\Sigma_{t|s} \equiv Var(x^f_t|{}_tX^f_{t-1}, X^p_s)$ for $s = t-1, t$. Under the normality and linearity assumptions. Equations (7) and (8) are computed as follows.

$$x^f_{t|t-1} = T_t\ {}_tx^f_{t-1},$$

$$\Sigma_{t|t-1} = R_t\Sigma_\eta R_t',$$

$$k_t = \Sigma_{t|t-1}Z_t'(Z_t\Sigma_{t|t-1}Z_t' + S_t\Sigma_\epsilon S_t')^{-1},$$

$$x^f_{t|t} = x^f_{t|t-1} + k_t(x^p_t - Z_t x^f_{t|t-1}),$$

$$\Sigma_{t|t} = \Sigma_{t|t-1} - k_t(Z_t\Sigma_{t|t-1}Z_t' + S_t\Sigma_\epsilon S_t')k_t',$$

for $t = 1, \cdots, T$. $x^f_{t|t-1}$ and $\Sigma^f_{t|t-1}$ in the first two equations yields the predictor (iv) when $x^f_{t-1|t-1}$ in equation (10) and $\Sigma_{t-1|t-1}$ in equation (11) are replaced by the most recently revised data ${}_tx^f_{t-1}$ and zero, respectively. Also, the predictor (v) is given by $x^f_{t|t}$ and $\Sigma_{t|t}$ when $x^f_{t-1|t-1} = {}_tx^f_{t-1}$ and $\Sigma_{t-1|t-1} = 0$ in the recursive algorithm (10) – (14).

[10] When the initial value is normally distributed, $x^f_{0|0} = \mu_0$ and $\Sigma_{0|0} = \Sigma_0$ are given, where $x^f_0 \sim N(\mu_0, \Sigma_0)$.

3.2 Nonlinear and/or Nonnormal Case

The Monte-Carlo integration filter used in this section was proposed by Tanizaki (1991,1993) and Tanizaki and Mariano (1994), where the importance sampling theory presented by Geweke (1989) and Shao (1989) is applied to density approximation and a recursive algorithm of the weight function represented as the ratio of two density functions is derived. In this section, we briefly discuss the nonlinear filter algorithm with the importance sampling.

(i) $\boldsymbol{E(x_t^f|X_{t-1}^p)}$ and (ii) $\boldsymbol{E(x_t^f|X_t^p)}$:

Define the weight function as $\omega_{i,t|s} = \dfrac{P(x_{i,t}^f|X_s^p)}{P_x(x_{i,t}^f)}$ for $s = t-1, t$, where $P_x(x_t^f)$ denotes the importance density and $x_{i,t}^f$, $i = 1,\cdots,n$, are the random draws generated from $P_x(x_t^f)$. Using the weight function, equations (3) and (4) can be rewritten as:

$$\omega_{i,t|t-1} = \begin{cases} \dfrac{P(x_{i,t}^f|x_{t-1}^f)}{P_x(x_{i,t}^f)}, & \text{if } t=1, \\[2ex] \dfrac{1}{n}\displaystyle\sum_{j=1}^{n} \dfrac{P(x_{i,t}^f|x_{j,t-1}^f)}{P_x(x_{i,t}^f)}\omega_{j,t-1|t-1}, & \text{otherwise}, \end{cases} \tag{19}$$

$$\omega_{i,t|t} = \frac{P(x_t^p|x_{i,t}^f)\omega_{i,t|t-1}}{\dfrac{1}{n}\displaystyle\sum_{j=1}^{n} P(x_t^p|x_{j,t}^f)\omega_{j,t|t-1}}. \tag{20}$$

where $t = 1,\cdots,T$ and $i = 1,\cdots,n$. Equations (19) and (20) show the filtering algorithm based on the weight functions. Since $P(x_t^p|X_{t-1}^p)$ represents the denominator of equation (20), the log-likelihood function using the weight function $\omega_{j,t|t-1}$ is given by:

$$\log L(\gamma,\theta;X_t^p) = \sum_{t=1}^{T} \log\Big(\frac{1}{n}\sum_{j=1}^{n} P(x_t^p|x_{j,t}^f)\omega_{j,t|t-1}\Big). \tag{21}$$

(iii) $\boldsymbol{E(x_t^f|X_T^p)}$:

Define the weight function as $\omega_{i,t|T} = \dfrac{P(x_{i,t}^f|X_T^p)}{P_x(x_{i,t}^f)}$. Equation (6) is approximated as:

$$\omega_{i,t-1|T} = \omega_{i,t-1|t-1}\frac{1}{n}\sum_{j=1}^{n} \frac{\omega_{j,t|T}P(x_{j,t}^f|x_{i,t-1}^f)}{\omega_{j,t|t-1}P_x(x_{j,t}^f)},$$

where $t = T, T-1, \cdots, 1$ and $i = 1,\cdots,n$.

(iv) $\boldsymbol{E(x_t^f|_t X_{t-1}^f, X_{t-1}^p)}$ and (v) $\boldsymbol{E(x_t|_t X_{t-1}^f, X_t^p)}$:

For the predictors (iv) and (v) based on the state-space model (1) and (2′), re-define the weight function as $\omega_{i,t|s} = \dfrac{P(x_{i,t}^f|_t X_{t-1}^f, X_s^p)}{P_x(x_{i,t}^f)}$ for $s = t-1, t$. By the importance sampling, equations (7) and (8) are represented by:

$$\omega_{i,t|t-1} = \frac{P(x_{i,t}^f|_t x_{t-1}^f)}{P_x(x_{i,t}^f)},$$

$$\omega_{i,t|t} = \frac{P(x_t^p|x_{i,t}^f)\omega_{i,t|t-1}}{\dfrac{1}{n}\displaystyle\sum_{j=1}^{n} P(x_t^p|x_{j,t}^f)\omega_{j,t|t-1}},$$

where $t = 1, \cdots, T$ and $i = 1, \cdots, n$.

Thus, all the algorithms for prediction have been derived using the weight functions. Each expectation and its variance are evaluated as follows:

$$E(x_t^f|\Omega_j) = \frac{1}{n}\sum_{i=1}^{n} x_{i,t}^f \omega_{i,t|s},$$
$$Var(x_t^f|\Omega_j) = \frac{1}{n}\sum_{i=1}^{n} \big(x_{i,t}^f - E(x_t^f|\Omega_j)\big)^2 \omega_{i,t|s},$$

for $s = t-1, t, T$ and $j = 1, \cdots, 5$, where Ω_j denotes the appropriate information set, i.e., $\Omega_1 = X_{t-1}^p$, $\Omega_2 = X_t^p$, $\Omega_3 = X_T^p$, $\Omega_4 = \{{}_tX_{t-1}^f, X_{t-1}^p\}$, and $\Omega_5 = \{{}_tX_{t-1}^f, X_t^p\}$, and the appropriate weight function (i.e., $\omega_{i,t|s}$) corresponding to the information set Ω_j has to be used for (i) – (v).

It is very important for the above Monte-Carlo integration procedure to note that computation errors become drastically large as t is large. In order to avoid accumulation of the computation errors, we need the weight function $\omega_{i,t|s}$ which satisfies $\frac{1}{n}\sum_{i=1}^{n}\omega_{i,t|s} = 1$ for $s = t-1, t, T$ and $t = 1, \cdots, T$, because integration of a density function is equal to one.

The advantages of using the Monte-Carlo integration filter are less computational burden and easier computer programming, especially in the higher dimensional cases of x_t^f, than the other distribution-based nonlinear filters proposed by Kitagawa (1987), Carlin, Polson and Stoffer (1992) and so on. On the other hand, the disadvantage is that precision of density approximation depends on the importance density $P_x(x_t^f)$. If $P_x(x_t^f)$ is chosen away from $P(x_t^f|X_{t-1}^p)$ and $P(x_t^f|X_t^p)$ in equations (3) and (4) or $P(x_t^f|{}_tX_{t-1}^f, X_{t-1}^p)$ and $P(x_t^f|{}_tX_{t-1}^f, X_t^p)$ in equations (7) and (8), the weight functions become unrealistic. Accordingly, the filtering estimates based on the weight functions are biased in such a case.[11]

Thus, in the similar context, the predictors (i) – (v) have been introduced. Considering that we do not really have the final data because they are always revised, we take the final data as unobservable variables and moreover estimate the parameters and the final data (the state variable) simultaneously, using the filtering or smoothing technique. (iv) and (v) are computed utilizing the past revised data ${}_tX_{t-1}^f$, since ${}_tX_{t-1}^f$ is available in addition to the preliminary data up to time $t-1$ or t, i.e., X_s^p for $s = t-1, t$. In the next section, a representative consumption function of U.S. is estimated as an application.

4 EMPIRICAL EXAMPLE

In this section, we compare the predictors of final data based on preliminary data and/or revised data. A consumption function using U.S. national accounts data is estimated to illustrate the revision process. The measurement and transition equations in the state-space model are derived as follows.

Measurement Equation: Since the preliminary data is related to the final data, simply we consider the measurement equation as the following linear or log-linear relationship between x_t^p and x_t^f:

$$x_t^p = \gamma_0 + \gamma_1 x_t^f + \epsilon_t, \qquad \epsilon_t \sim N\big(0, \gamma_2^2 (x_t^f)^2\big), \tag{22}$$
$$\log(x_t^p) = \gamma_0 + \gamma_1 \log(x_t^f) + \epsilon_t, \qquad \epsilon_t \sim N\big(0, \gamma_2^2 (\log(x_t^f))^2\big), \tag{23}$$

where it is assumed for the error ϵ_t to be heteroskedastic, i.e., the variance of ϵ_t depends on the state vector. γ_0, γ_1 and γ_2 are the unknown parameters.[12]

For each measurement equation of (22) and (23), we consider the following three types (a) – (c).

(a) $\gamma_0 = 0$ and $\gamma_1 = 1$ are assumed but γ_2 is the unknown parameter to be estimated.

(b) γ_0, γ_1 and γ_2 are the unknown parameters to be estimated.

[11] For choice of the importance density, see APPENDIX.

[12] Note that the unknown parameter vector γ in equation (1) is denoted by $\gamma = \{\gamma_0, \gamma_1, \gamma_2\}$ in equations (22) and (23).

(c) γ_0, γ_1 and γ_2 are *a priori* estimated by weighted ordinary least squares estimation (WOLS) using the most recently revised data (the real per capita consumption data taken from the ERP published in 1994) for x_t^f and the preliminary data for x_t^p.

The WOLS results are as follows:

$$x_t^p = \underset{(59.378)}{372.508} + \underset{(0.007663)}{0.921857}\, x_t^f, \qquad se = 0.01296, \quad DW = 0.2677.$$

$$\log(x_t^p) = \underset{(0.072033)}{0.373297} + \underset{(0.007991)}{0.955054}\, \log(x_t^f), \qquad se = 0.001572, \quad DW = 0.2426.$$

The values in the parentheses represent the standard errors of the coefficients. se and DW denote γ_2 and Durbin-Watson ratio.

The estimation period is 1947 – 1989 for both equations.[13] For x_t^f, the most recently revised consumption data from 1947 to 1989 (i.e., ${}_s x_t^f$ for $t = 1947, \cdots, 1989$ and $s = 1993$) are taken from the ERP published in 1994. That is, equations (22) and (23) are estimated replacing x_t^f by ${}_s x_t^f$, $s = 1993$.

Transition Equation: The consumption function derived from a representative utility maximization problem is taken as the transition equation. The Euler equation can be nonlinear in parameters and/or error terms, depending on the functional form of the underlying utility function. Accordingly, the nonlinear filtering techniques introduced in Section 3.2 can be applicable and useful.

Consider the problem of choosing a consumption sequence of the representative agent which maximizes the expected utility:

$$E_t\Big(\sum_t \beta^t u(x_t, \mu_t)\Big), \tag{24}$$

subject to $A_{t+1} = R_t(A_t + y_t - x_t)$, where x_t is consumption at time t and A_0 is given. $E_t(\cdot)$ denotes the mathematical expectation, given information known at t. The representative utility function $u(x_t, \mu_t)$ is twice continuously differentiable, bounded, increasing in consumption x_t, and concave in x_t. A_{t+1} is the stock of assets at the beginning of time $t+1$, y_t is noncapital or labor income at t, and R_t is the gross rate of return on savings between t and $t+1$. x_t is assumed to be stochastic while R_t is nonstochastic for simplicity. β is the discount rate.[14] μ_t represents an exogenous part of consumption which does not depends on the permanent income hypothesis proposed by Hall (1978).

Maximizing the expected utility (24) with respect to the consumption sequence $\{x_t\}$, we obtain the Euler equation: $E_t\big(\beta R_t u'(x_{t+1}, \mu_{t+1})\big) = u'(x_t, \mu_t)$. Here, we take the following utility functions:

$$u(x_t, \mu_t) = -\frac{1}{2}\big(\overline{x} - (x_t - \mu_t)\big)^2, \text{ and } \beta R_t = 1, \tag{25}$$

$$u(x_t, \mu_t) = -\frac{1}{\alpha} e^{-\alpha(x_t - \mu_t)}, \tag{26}$$

$$u(x_t, \mu_t) = \frac{(x_t/\mu_t)^{1-\alpha} - 1}{1-\alpha}, \qquad \text{where } \alpha \text{ indicates a measure of risk aversion.} \tag{27}$$

In the above utility functions (25) – (27), it is assumed that a part of consumption (i.e., $x_t - \mu_t$ or x_t/μ_t) follows the permanent income hypothesis because it is well known that variables other than lagged consumption appear to play a significant role in the determination of current consumption (see Diebold

[13]Consumption data from 1947 to 1989 are not revised in the ERP published in 1994. Therefore, we use the data from 1947 to 1989. See Table 1.

[14]It might be plausible in empirical studies that the discount rate β is less than one. However, Kocherlakota (1990) showed that well-behaved competitive equilibria with positive interest rates may exist in infinite horizon growth economies even though individuals have discount factors larger than one, which implies that when per capita consumption is growing over time, it is possible for equilibria to exist in representative consumer endowment economies even though $\beta > 1$. Therefore, we do not have to pay much attention to the discount rate greater than one.

Table 2: Unknown Parameters to be Estimated by MLE

ME \ TE		(28)	(29)	(30)
(22)	(a)	γ_2,δ,σ	$\gamma_2,\alpha,\beta,\delta,\sigma$	$\gamma_2,\alpha,\beta,\delta,\sigma$
	(b)	$\gamma_0,\gamma_1,\gamma_2,\delta,\sigma$	$\gamma_0,\gamma_1,\gamma_2,\alpha,\beta,\delta,\sigma$	$\gamma_0,\gamma_1,\gamma_2,\alpha,\beta,\delta,\sigma$
	(c)	δ,σ	$\alpha,\beta,\delta,\sigma$	$\alpha,\beta,\delta,\sigma$
(23)	(a)	γ_2,δ,σ	$\gamma_2,\alpha,\beta,\delta,\sigma$	$\gamma_2,\alpha,\beta,\delta,\sigma$
	(b)	$\gamma_0,\gamma_1,\gamma_2,\delta,\sigma$	$\gamma_0,\gamma_1,\gamma_2,\alpha,\beta,\delta,\sigma$	$\gamma_0,\gamma_1,\gamma_2,\alpha,\beta,\delta,\sigma$
	(c)	δ,σ	$\alpha,\beta,\delta,\sigma$	$\alpha,\beta,\delta,\sigma$

and Nerlove (1989)). μ_t is a part of consumption which depends on the other variable such as time trend or income.[15]

Additional assumptions are made as follows: For (25), $x_t - x_{t-1}$ is assumed to be normal with mean zero and variance σ^2 and $\mu_t = \delta y_t$ is assumed. For (26), $x_t - x_{t-1}$ is assumed to be normally distributed with variance σ^2 and $\mu_t = \delta y_t$ is assumed. For (27), $\log(x_t) - \log(x_{t-1})$ is assumed to be normally distributed with variance σ^2 and $\log(\mu_t) = \delta\log(y_t)$ is assumed. Moreover, x_t in (25) – (27) should be replaced by the final consumption data x_t^f, because as discussed in Section 2 it might be plausible to consider that the variables included in the theoretical model do not have the measurement errors.

From (25) – (27), the consumption functions are written as:

$$x_t^f - x_{t-1}^f = \delta(y_t - y_{t-1}) + \eta_t, \tag{28}$$

$$x_t^f - x_{t-1}^f = \frac{1}{\alpha}\log(\beta R_{t-1}) + \frac{1}{2}\alpha\sigma^2 + \delta(y_t - y_{t-1}) + \eta_t, \tag{29}$$

$$\log\Big(\frac{x_t^f}{x_{t-1}^f}\Big) = \frac{1}{\alpha}\log(\beta R_{t-1}) + \frac{1}{2}\alpha\sigma^2 + \delta\big(1 - \frac{1}{\alpha}\big)\log\Big(\frac{y_t}{y_{t-1}}\Big) + \eta_t, \tag{30}$$

where x_t^f and y_t denote per capita final consumption and per capita income and η_t is normally distributed with mean zero and variance σ^2. Thus, the Euler equations (28) – (30) are used for the transition equation (2) in the state-space model, where the underlying assumption is that current consumption depends on not only lagged consumption but also variables other than the lagged consumption.[16]

Estimation Results: Thus, we have equations (22), (23) and (a) – (c) for the measurement equation and equations (28) – (30) for the transition equation. The log-likelihood function (15) or (21) is maximized with respect to the unknown parameters shown in Table 2 by a simple grid search.[17] Two-step procedure is taken for estimation, which is as follows.

1. The first step is to perform the extended Kalman filter, which is the nonlinear filter based on the first-order Taylor series expansion. The measurement equation (22) or (23) and the transition

[15] The maximization procedure of equation (24) is equivalent to maximizing the following expected utility function with respect to x_t^*:

$$E_t\Big(\sum_t \beta^t u(x_t^*)\Big),$$

subject to $A_{t+1} = R_t(A_t + y_t - x_t)$, where $x_t = x_t^* + \mu_t$ or $x_t = x_t^*\mu_t$ (i.e., $\log(x_t) = \log(x_t^*) + \log(\mu_t)$). x_t^* indicates the consumption which depends on the permanent income hypothesis.

[16] Note that the unknown parameter vector θ in equation (2) is given by $\theta = \{\delta, \sigma\}$ for (28), and $\theta = \{\alpha, \beta, \delta, \sigma\}$ for (29) and (30).

[17] Six types of measurement equations (i.e., (22), (23) and (a) – (c)) and three types of transition equations (i.e., (28) – (30)) are estimated (6×3 state-space models are considered). For each state-space model, the unknown parameters to be estimated are summarized in Table 2. Note that ME and TE in Tables 2 – 6 denotes the measurement equation and the transition equation.

equation (28), (29) or (30) are linearized with respect to (x_t^f, ϵ_t) and (x_{t-1}^f, η_t). Then, the linearized nonlinear equations are applied to the algorithms shown in Section 3.1. Note that the transition equations (28) and (29) are already linear but the measurement equations (22) and (23) are not (recall that variance of ϵ_t is assumed to be $\gamma_2^2(x_t^f)^2$ or $\gamma_2^2\big(\log(x_t^f)\big)^2$ which is nonlinear in x_t^f). The log-likelihood function (15) is maximized with respect to the unknown parameters by a simple grid search.

For (i) and (ii), based on the extended Kalman filter, The unknown parameters are estimated together with $E(x_t^f|\Omega_i)$ and $Var(x_t^f|\Omega_i)$, $i = 1, 2$. The estimates of $\{\gamma_0,\ \gamma_1,\ \gamma_2,\ \delta,\ \sigma,\ \alpha,\ \beta\}$ by the extended Kalman filter are in Table 3. For (iii) – (v), given the parameter estimates, $E(x_t^f|\Omega_i)$ and $Var(x_t^f|\Omega_i)$, $i = 3, 4, 5$, are obtained by applying the linearized nonlinear functions to the appropriate linear algorithms in Section 3.1.

2. In the second step, based on the final consumption estimates obtained from the extended Kalman filter (i.e., Step 1), the Monte-Carlo integration procedure shown in Section 3.2 is implemented to obtain more precise estimate of the state variable than the extended Kalman filter. There, the importance density $P_x(x_t^f)$ for (i) and (ii) is taken as:

$$P_x(x_t^f) = \frac{1}{2}N(\mu_{t|1}^*, 4\Sigma_{t|1}^*) + \frac{1}{2}N(\mu_{t|2}^*, 4\Sigma_{t|2}^*).$$

$P_x(x_t^f)$ is a bimodal distribution based on the two normal densities.[18] $\mu_{t|1}^*$ and $\mu_{t|2}^*$ are the estimate of $E(x_t^f|X_{t-1}^p)$ and that of $E(x_t^f|X_t^p)$, which are obtained by the extended Kalman filter in Step 1. $\Sigma_{t|1}^*$ and $\Sigma_{t|2}^*$ are also the estimate of $Var(x_t^f|X_{t-1}^p)$ and that of $Var(x_t^f|X_t^p)$ taken from Step 1. The parameter estimates by the extended Kalman filter (i.e., Table 3) are utilized for the initial parameter values. Moreover, we set the number of random draws as $n = 200$ and maximize the likelihood function (21) by a simple grid search in order to update the parameter estimates of $\{\gamma_0,\ \gamma_1,\ \gamma_2,\ \delta,\ \sigma,\ \alpha,\ \beta\}$ to Table 4. Thus, for the predictors (i) and (ii), $E(x_t^f|\Omega_i)$ and $Var(x_t^f|\Omega_i)$, $i = 1, 2$, are re-computed utilizing the nonlinear algorithms in Section 3.2.

For (iv) and (v), the same importance density is taken, but $\mu_{t|1}^*$ and $\mu_{t|2}^*$ are the estimates of the predictors $E(x_t^f|{}_tX_{t-1}^f, X_{t-1}^p)$ and $E(x_t^f|{}_tX_{t-1}^f, X_t^p)$, which are obtained from the extended Kalman filter (i.e., Step 1), and $\Sigma_{t|1}^*$ and $\Sigma_{t|2}^*$ are also the estimate of $Var(x_t^f|{}_tX_{t-1}^f, X_{t-1}^p)$ and that of $Var(x_t^f|{}_tX_{t-1}^f, X_t^p)$ from Step 1.

Thus, the Monte-Carlo integration filter proposed in Section 3.2 takes the two-step procedure.

The measurement equation (22) or (23) and the transition equation (28), (29), or (30) are estimated in Tables 3 and 4.[19] The estimates in Table 3 are very close to those in Table 4. Given the estimates

[18] The importance density $P_x(x_t^f)$ has to be used for approximation of the two densities $P(x_t^f|X_s^p)$, $s = t-1, t$. $x_{t|s}^f$ and $\Sigma_{t|s}$, $s = t-1, t$, computed from the extended Kalman filter are biased but not too far from the true values. From APPENDIX, we should choose the importance density which is broadly distributed, compared with the original density. Therefore, the bimodal density utilizing the extended Kalman filter is taken for the importance density, i.e., $P_x(x_t^f) = \frac{1}{2}N(\mu_{t|1}^*, 4\Sigma_{t|1}^*) + \frac{1}{2}N(\mu_{t|2}^*, 4\Sigma_{t|2}^*)$. See APPENDIX for precision of Monte-Carlo integration and choice of this importance density.

[19] All the data used for estimation are as follows.

The preliminary data x_t^p is the per capita real preliminary consumption data (dollars). In Table 1, the data with the superscript p represents Personal Consumption Expenditures (billions of dollars), which is the nominal preliminary data. In order to obtain x_t^p, the nominal preliminary data shown in Table 1 is converted to the per capita real data divided by L_t and P_t, where L_t and P_t denote U.S. population (billions of people) and Implicit Price Deflator of Personal Consumption Expenditures (1982=1.00). Both L_t and P_t are taken from the ERP published in 1994.

The interest rate R_t included in the Euler equations (29) and (30) is computed as $R_t = (1 + r_t/100)P_t/P_{t+1}$, where r_t represents Moody's Corporate Bonds, Aaa (percent) taken from the ERP published in 1994. Moreover, we use per capita real Disposable Income for y_t, which is taken from the ERP published in 1994. Y_t in Table 1 shows nominal Disposable Income, which is converted to the per capita real data divided by L_t and P_t (i.e., $y_t = Y_t/L_tP_t$).

In this paper, because we focus on consumption, it is assumed that P_t, L_t, y_t and r_t are nonstochastic, which implies that the data P_t, L_t, y_t and r_t are not revised (i.e., that the final data only are reported). This is true for r_t but not for L_t, y_t and P_t. y_t, L_t and P_t are also revised every year for the first few years and thereafter they are revised less frequently. Here, however, we assume for simplicity of discussion that L_t, y_t and P_t do not have the measurement errors.

Table 3: Estimation of Unknown Parameters: Extended Kalman Filter

TE	ME		α	β	δ	σ	γ_0	γ_1	γ_2	$\log L(\gamma,\theta;X_t^p)$ (15)
(28)	(22)	(a)	—	—	0.772	116	0	1	0.00564	−296.93
		(b)	—	—	0.790	94	86	1.046	0.00658	−292.44
		(c)	—	—	0.872	117	372.5	0.9219	0.01296	−303.76
	(23)	(a)	—	—	0.771	116	0	1	0.000615	−296.81
		(b)	—	—	0.768	90	−0.0268	1.010	0.000691	−292.09
		(c)	—	—	0.870	121	0.3733	0.9551	0.001572	−305.63
(29)	(22)	(a)	0.000727	1.031	0.498	99	0	1	0.00524	−290.40
		(b)	0.000497	0.996	0.611	63	108	1.062	0.00695	−282.19
		(c)	0.000694	1.041	0.483	85	372.5	0.9219	0.01296	−297.94
	(23)	(a)	0.000730	1.031	0.499	99	0	1	0.000571	−290.28
		(b)	0.000511	0.995	0.600	60	0.00994	1.008	0.000720	−281.80
		(c)	0.000685	1.043	0.463	88	0.3733	0.9551	0.001572	−299.87
(30)	(22)	(a)	9.45	1.077	0.476	0.0126	0	1	0.00300	−288.07
		(b)	4.43	1.003	1.019	0.0104	1200	0.842	0.00467	−278.67
		(c)	10.17	1.096	0.509	0.0062	372.5	0.9219	0.01296	−294.63
	(23)	(a)	9.54	1.078	0.475	0.0128	0	1	0.000286	−288.08
		(b)	4.62	1.000	0.846	0.0091	0.00421	1.007	0.000507	−279.42
		(c)	10.09	1.102	0.474	0.0065	0.3733	0.9551	0.001572	−296.51

Table 4: Estimation of Unknown Parameters: Monte-Carlo Integration Filter

TE	ME		α	β	δ	σ	γ_0	γ_1	γ_2	$\log L(\gamma,\theta;X_t^p)$ (21)
(28)	(22)	(a)	—	—	0.775	116	0	1	0.00567	−296.90
		(b)	—	—	0.807	95	194	1.024	0.00655	−292.31
		(c)	—	—	0.879	116	372.5	0.9219	0.01296	−303.66
	(23)	(a)	—	—	0.775	115	0	1	0.000623	−296.78
		(b)	—	—	0.771	90	−0.0214	1.009	0.000694	−292.07
		(c)	—	—	0.879	120	0.3733	0.9551	0.001572	−305.50
(29)	(22)	(a)	0.000737	1.032	0.500	98	0	1	0.00534	−290.35
		(b)	0.000487	0.996	0.627	64	245	1.036	0.00693	−282.06
		(c)	0.000703	1.042	0.487	84	372.5	0.9219	0.01296	−297.79
	(23)	(a)	0.000739	1.032	0.500	98	0	1	0.000583	−290.23
		(b)	0.000516	0.995	0.602	60	0.00900	1.008	0.000723	−281.81
		(c)	0.000693	1.044	0.468	86	0.3733	0.9551	0.001572	−299.68
(30)	(22)	(a)	9.58	1.077	0.486	0.0125	0	1	0.00316	−287.92
		(b)	4.56	1.005	1.012	0.0104	1290	0.822	0.00473	−278.47
		(c)	10.30	1.097	0.513	0.0062	372.5	0.9219	0.01296	−294.53
	(23)	(a)	9.66	1.078	0.484	0.0125	0	1	0.000345	−287.93
		(b)	4.83	1.003	0.816	0.0091	0.00007	1.007	0.000513	−279.39
		(c)	10.28	1.103	0.484	0.0063	0.3733	0.9551	0.001572	−296.29

of (γ, θ) in Table 4, the predictors (i) – (v) are compared in Tables 5 and 6, which are computed based on the nonlinear procedures shown in Section 3.2. Since the true final data is unobservable for any t, the estimated final data and L-th revised data, $L = 1, 2, \cdots, 10$, are compared in the criteria of mean absolute percent error ($MAPE^{(L)}$) and weighted root mean square error ($WRMSE^{(L)}$).[20] As mentioned above, each information set is represented by $\Omega_1 = X^p_{t-1}$, $\Omega_2 = X^p_t$, $\Omega_3 = X^p_T$, $\Omega_4 = \{{}_tX^f_{t-1}, X^p_{t-1}\}$, and $\Omega_5 = \{{}_tX^f_{t-1}, X^p_t\}$. The criteria $MAPE^{(L)}_j$ and $WRMSE^{(L)}_{ij}$, $i, j = 1, \cdots, 5$, are defined as follows:

$$MAPE^{(L)}_j = 100 \times \frac{1}{T-L} \sum_{t=1}^{T-L} \left| {}_{t+L}x^f_t - E(x^f_t|\Omega_j) / {}_{t+L}x^f_t \right|,$$

$$WRMSE^{(L)}_{ij} = \sqrt{\frac{1}{T-L} \sum_{t=1}^{T-L} \left({}_{t+L}x^f_t - E(x^f_t|\Omega_j) \right)^2 / Var(x^f_t|\Omega_i)},$$

where ${}_{t+L}x^f_t$ implies the L-th revised data of time t. We compare precision of the predictors (i) – (v) by the $MAPE$ and $WRMSE$ criteria. The superscript L in $MAPE^{(L)}_j$ and $WRMSE^{(L)}_{ij}$ implies comparison between the L-th revised data and the estimated final data, while the subscript j indicates precision of the predictors (i) – (v). Furthermore, the subscript i in $WRMSE^{(L)}_{ij}$ represents size of each variance for the predictors (i) – (v). Thus, $MAPE^{(L)}$ and $WRMSE^{(L)}$ in Tables 5 and 6 denote a 1×5 vector and a 5×5 matrix, i.e., $MAPE^{(L)} = \{MAPE^{(L)}_j\}$ and $WRMSE^{(L)} = \{WRMSE^{(L)}_{ij}\}$ for $L = 1,\ 2,\ 3,\ 6,\ 10$. In Tables 5 and 6, each column (i.e., the subscript j) in $MAPE^{(L)}_j$ and $WRMSE^{(L)}_{ij}$ represents precision of the predictors, while each row (i.e., the subscript i) in $WRMSE^{(L)}_{ij}$ indicates normalization by $Var(x^f_t|\Omega_i)$. For each of Tables 5 and 6, several important issues to be analyzed are in Sections 4.1 – 4.5.

4.1 $MAPE^{(L)}_j$ for Functional Forms (See Table 5):

From the $MAPE^{(L)}_j$ criterion given j and L, we examine which functional form we should choose for the measurement equations (22).(a) – (23).(c) and the transition equations (28) – (30)[21]. For the true final data, we should choose the predictor (i.e., $j = 1, \cdots, 5$) such that $MAPE^{(L)}_j$ is the smallest when L is large, because it might be appropriate to consider that the L-th revised data approaches the final data as L increases. When $L = 10$ is taken, the smallest $MAPE^{(L)}_j$ is 1.64 – 1.67 for (v) of model (a), 1.69 – 1.84 for (iv) of model (b), and 1.43 – 1.56 for (v) of model (c). For both (22) and (23), model (b) has a poor performance because $MAPE^{(L)}_j$ is extremely large except for $j = 4$. Accordingly, hereafter we do not examine model (b) for the cases of both (22) and (23).

For economic forecasting, the first (or second) revised data is sometimes important. In this case, we should choose the predictor (i.e., $j = 1, \cdots, 5$) with the smallest $MAPE^{(L)}_j$ for small L. Taking an example of the case: $j = 1$ and $L = 1$, we have 1.34 for (22).(a) and (28), 1.11 for (22).(a) and (29), 1.13 for (22).(a) and (30), 5.32 for (22).(b) and (28), 6.92 for (22).(b) and (29), 4.82 for (22).(b) and (30), 2.73 for (22).(c) and (28), 3.10 for (22).(c) and (29), and 2.97 for (22).(c) and (30). For both measurement equations (22) and (23), model (a) gives us the smallest $MAPE$. For the transition equation, (29) is the smallest in almost cases.

[20]The reason why the *weighted* root mean square error (i.e., $WRMSE^{(L)}$) is taken for comparison is because ${}_{t+L}x^f_t - E(x^f_t|\Omega_j)$, $j = 1, \cdots, 5$, are heteroskedastic over time t. Therefore, we should take the root mean square error *weighted* at each standard error as the criterion of prediction precision.

[21]Note that $MAPE^{(L)}_j$ should be compared for the following cases: Section 4.1: $MAPE^{(L)}_j$ for Functional Forms and Section 4.2: $MAPE^{(L)}_j$ vs $MAPE^{(L)}_{j'}$ for $j \neq j'$.

In Section 4.1, we consider which functional form is the best underlying assumption, while in Section 4.2 we examine which predictor is the best.

Given j, $MAPE^{(L)}_j$ cannot be compared for different L, because the denominator ${}_{t+L}x^f_t$ depends on L from the definition of $MAPE^{(L)}_j$.

Table 5: Precision of Final Data: Mean Absolute Percent Error (i.e., $MAPE_j^{(L)}$)

ME		TE	(28)					(29)					(30)				
		L	(i)	(ii)	(iii)	(iv)	(v)	(i)	(ii)	(iii)	(iv)	(v)	(i)	(ii)	(iii)	(iv)	(v)
		j	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
(22)	(a)	1	1.34	0.76	0.69	1.01	0.70	1.11	0.73	0.68	0.97	0.67	1.13	0.72	0.71	0.99	0.70
		2	1.57	1.01	0.95	1.17	0.92	1.36	0.97	0.95	1.09	0.88	1.37	0.95	0.95	1.10	0.92
		3	1.71	1.18	1.13	1.34	1.12	1.45	1.15	1.13	1.20	1.08	1.46	1.15	1.14	1.19	1.12
		6	1.96	1.48	1.43	1.60	1.42	1.66	1.45	1.44	1.40	1.39	1.69	1.46	1.46	1.41	1.44
		10	1.99	1.70	1.70	1.84	1.66	1.85	1.69	1.71	1.73	1.64	1.90	1.69	1.70	1.74	1.67
	(b)	1	5.32	5.33	5.27	1.02	4.13	6.92	6.99	6.99	0.94	4.19	4.82	4.86	4.87	1.02	3.77
		2	5.62	5.63	5.58	1.18	4.45	7.24	7.30	7.30	1.05	4.53	4.54	4.63	4.63	1.01	3.56
		3	5.76	5.78	5.74	1.35	4.63	7.41	7.47	7.46	1.16	4.73	4.33	4.42	4.43	1.06	3.36
		6	6.03	6.09	6.06	1.60	5.00	7.71	7.79	7.78	1.45	5.15	3.71	3.82	3.84	1.34	2.84
		10	6.17	6.34	6.37	1.84	5.34	7.98	8.09	8.12	1.81	5.59	3.17	3.12	3.12	1.69	2.30
	(c)	1	2.73	2.67	2.85	1.05	1.28	3.10	2.86	2.87	1.02	1.16	2.97	2.91	2.92	0.99	1.03
		2	2.48	2.38	2.58	1.21	1.15	2.76	2.55	2.54	1.11	1.05	2.61	2.57	2.61	1.08	0.97
		3	2.34	2.18	2.39	1.37	1.15	2.55	2.31	2.35	1.19	1.06	2.40	2.36	2.44	1.16	0.98
		6	2.20	1.83	2.00	1.61	1.31	2.31	1.96	2.00	1.39	1.18	2.08	1.97	2.11	1.36	1.14
		10	2.18	1.74	1.69	1.87	1.56	2.35	1.86	1.74	1.73	1.51	2.10	1.88	1.89	1.69	1.43
(23)	(a)	1	1.34	0.77	0.69	1.01	0.70	1.11	0.73	0.68	0.97	0.68	1.13	0.72	0.71	0.99	0.70
		2	1.57	1.01	0.95	1.17	0.92	1.36	0.98	0.95	1.09	0.88	1.38	0.95	0.95	1.10	0.92
		3	1.71	1.18	1.13	1.34	1.12	1.45	1.15	1.13	1.20	1.08	1.46	1.15	1.14	1.19	1.12
		6	1.96	1.48	1.43	1.60	1.42	1.66	1.46	1.44	1.40	1.39	1.70	1.46	1.46	1.41	1.44
		10	1.99	1.70	1.70	1.84	1.66	1.86	1.69	1.71	1.73	1.65	1.90	1.69	1.70	1.74	1.67
	(b)	1	6.36	6.37	6.31	1.01	4.83	8.24	8.30	8.29	0.93	4.84	6.62	6.65	6.65	0.94	5.36
		2	6.64	6.64	6.59	1.17	5.13	8.52	8.57	8.56	1.07	5.17	6.91	6.92	6.92	1.01	5.64
		3	6.74	6.76	6.71	1.33	5.30	8.65	8.70	8.69	1.19	5.38	7.05	7.05	7.06	1.10	5.79
		6	6.91	6.97	6.93	1.60	5.63	8.84	8.91	8.89	1.49	5.78	7.30	7.29	7.28	1.40	6.04
		10	6.93	7.09	7.11	1.84	5.90	8.97	9.05	9.08	1.84	6.17	7.44	7.45	7.46	1.77	6.24
	(c)	1	2.73	2.66	2.86	1.05	1.25	3.10	2.85	2.86	1.03	1.13	2.98	2.91	2.92	1.01	1.02
		2	2.46	2.37	2.57	1.21	1.13	2.77	2.54	2.54	1.12	1.04	2.60	2.55	2.59	1.09	0.97
		3	2.31	2.15	2.37	1.37	1.13	2.54	2.30	2.32	1.20	1.05	2.37	2.32	2.40	1.17	1.00
		6	2.18	1.78	1.96	1.61	1.30	2.30	1.95	1.97	1.40	1.19	2.05	1.93	2.07	1.36	1.16
		10	2.19	1.71	1.65	1.87	1.55	2.38	1.86	1.70	1.75	1.52	2.09	1.86	1.85	1.70	1.46

Thus, we should take model (a) when L is small, but model (a) or (c) when L is large. In the case of (22).(a) and (23).(a), equation (26) is the best assumption for small L. However, in the case of (22).(c) and (22).(c), equation (25) for (i) – (iii) and equation (27) for (iv) and (v) perform better for large L.

4.2 $MAPE_j^{(L)}$ vs $MAPE_{j'}^{(L)}$ for $j \neq j'$ (See Table 5):

Given L, $MAPE_j^{(L)}$ is compared for $j = 1, \cdots, 5$. In this section, from the $MAPE$ criterion we consider which is the best predictor of (i) – (v). In Table 5, predictor (v) shows the best performance for models (a) and (c), but not too different from predictors (ii) and (iii). Note that (iii) is better than (ii) because (iii) utilizes more information set than (ii). In the case of model (a), for almost all L, $j = 5$ gives the smallest value. For (c), (iv) is the best predictor when $L = 1$ but (v) is the best otherwise. Thus, the predictor (v) shows significant improvement in economic forecasting, compared with the other predictors. One of the reasons why (v) is the best predictor is because the information set for (v) is larger than any other predictor.

4.3 $WRMSE_{ij}^{(L)}$ vs $WRMSE_{ij'}^{(L)}$ for $j \neq j'$ (See Table 6):

Given i and L, $WRMSE_{ij}^{(L)}$ are compared for $j = 1, \cdots, 5$.[22] From the $WRMSE$ criterion, we consider which is the best predictor of (i) – (v). For models (a) and (c) in Table 6, (v) shows the best performance of the predictors ((iii) is sometimes the best for (a)), while the worst predictor is given by (i). For example, taking the case: (22).(a), (28), $i = 2$ and $L = 6$ in Table 6, $WRMSE_{ij}^{(L)}$ is 5.10 for $j = 1$, 3.80 for $j = 2$, 3.64 for $j = 3$, 4.27 for $j = 4$ and 3.64 for $j = 5$, respectively. In almost all the cases of (a) and (c), $j = 5$ gives the smallest value of all $j = 1, \cdots, 5$ for any i and any L. For models (a) and (c), (v) is close to (ii) and (iii), compared with (i) and (iv).

4.4 $WRMSE_{ij}^{(L)}$ vs $WRMSE_{ij}^{(L')}$ for $L \neq L'$ (See Table 6):

Given i and j, $WRMSE_{ij}^{(L)}$ is compared for $L = 1, 2, \cdots, 10$, because the mean square errors are divided by $Var(x_t^f|\Omega_i)$ independent of L. For the predictors (i), (ii) and (iii) in model (c), $WRMSE_{ij}^{(L)}$ decreases as L increases. However, $WRMSE_{ij}^{(L)}$ increases as L is large for (iv) in model (c), while it is almost constant over L for (v) in model (c) (see Table 6).

For all the predictors in model (a), $WRMSE_{ij}^{(L)}$ increases as L is large. Therefore, the final data estimated by model (a) is close to the L-th revised data for small L.

Since it is appropriate to consider that the revised data approaches the final data every time the data is revised, it is better for the true final data to choose the predictor such that $WRMSE_{ij}^{(L)}$ is small when L is large. For $j = 1, 2, 3$ of model (c), $WRMSE_{ij}^{(L)}$ decreases as L increases, which implies that the final data estimated using (c) is close to the true final data. For model (c), the best predictor of (i) – (iii) is (iii). Therefore, when we need to know the true final data, we should use the predictor (iii) of model (c).

Thus, it is appropriate to use model (a) when we need the true final data and model (c) when we want to know the first (or second) revised data (i.e., the L-th revised data for small L).

4.5 $WRMSE_{ij}^{(L)}$ vs $WRMSE_{i'j}^{(L)}$ for $i \neq i'$ (See Table 6):

Given j and L, $WRMSE_{ij}^{(L)}$ is compared for $i = 1, \cdots, 5$, which implies comparing variance for each predictor. From the definition of $WRMSE_{ij}^{(L)}$, $WRMSE_{ij}^{(L)}$ becomes large if variance of the predictor is

[22] Note that $WRMSE_{ij}^{(L)}$ should be compared for the following cases: Section 4.3: $WRMSE_{ij}^{(L)}$ vs $WRMSE_{ij'}^{(L)}$ for $j \neq j'$, Section 4.4: $WRMSE_{ij}^{(L)}$ vs $WRMSE_{ij}^{(L')}$ for $L \neq L'$ and Section 4.5: $WRMSE_{ij}^{(L)}$ vs $WRMSE_{i'j}^{(L)}$ for $i \neq i'$.

In Section 4.3, precision of each predictor is compared. In Section 4.4, we discuss comparison between the L-th revised data and the estimated final data. We consider which predictor has the smallest variance in Section 4.5.

Given i, j and L, it is meaningless to compare $WRMSE_{ij}^{(L)}$ for functional forms in Table 6. $Var(x_t^f|\Omega_i)$ in the denominator of $WRMSE_{ij}^{(L)}$ depends on the functional forms.

Table 6: Precision of Final Data: Weighted Root Mean Square Error (i.e., $WRMSE_{ij}^{(L)}$)

ME	L	TE	(28)					(29)					(30)				
		$i\backslash j$	(i) 1	(ii) 2	(iii) 3	(iv) 4	(v) 5	(i) 1	(ii) 2	(iii) 3	(iv) 4	(v) 5	(i) 1	(ii) 2	(iii) 3	(iv) 4	(v) 5
(22) (a)	1	1	1.29	0.79	0.67	0.92	0.73	1.28	0.88	0.78	1.02	0.82	1.15	0.75	0.74	1.02	0.74
		2	3.40	1.95	1.70	2.74	1.81	3.15	2.02	1.84	2.79	1.87	4.76	3.18	3.11	4.23	3.12
		3	3.62	2.09	1.81	2.89	1.94	3.39	2.19	1.99	2.96	2.03	4.90	3.27	3.19	4.35	3.20
		4	1.41	0.88	0.73	0.99	0.81	1.41	0.99	0.87	1.10	0.91	1.17	0.77	0.76	1.04	0.76
		5	3.43	1.98	1.72	2.76	1.84	3.20	2.06	1.87	2.81	1.91	4.77	3.19	3.11	4.23	3.12
	2	1	1.61	1.04	0.95	1.17	0.96	1.58	1.16	1.10	1.20	1.05	1.43	0.98	0.97	1.19	0.96
		2	4.19	2.56	2.39	3.37	2.39	3.86	2.66	2.56	3.24	2.45	5.88	4.10	4.07	4.89	4.01
		3	4.47	2.75	2.57	3.56	2.56	4.16	2.88	2.78	3.45	2.65	6.04	4.21	4.18	5.03	4.12
		4	1.76	1.15	1.05	1.26	1.06	1.75	1.30	1.24	1.30	1.18	1.45	1.00	0.99	1.21	0.98
		5	4.23	2.60	2.43	3.40	2.42	3.91	2.71	2.61	3.27	2.50	5.88	4.10	4.07	4.90	4.01
	3	1	1.73	1.24	1.15	1.33	1.17	1.69	1.38	1.33	1.30	1.29	1.49	1.16	1.15	1.22	1.14
		2	4.40	3.02	2.84	3.61	2.86	4.01	3.13	3.04	3.36	2.93	6.15	4.88	4.85	5.06	4.80
		3	4.71	3.25	3.05	3.84	3.08	4.35	3.41	3.31	3.59	3.20	6.32	5.01	4.98	5.20	4.93
		4	1.90	1.37	1.27	1.44	1.29	1.89	1.55	1.50	1.42	1.45	1.51	1.19	1.18	1.25	1.17
		5	4.45	3.07	2.88	3.65	2.91	4.08	3.19	3.10	3.40	3.00	6.15	4.88	4.85	5.06	4.80
	6	1	2.03	1.56	1.48	1.62	1.49	2.05	1.76	1.72	1.62	1.66	1.76	1.47	1.47	1.44	1.45
		2	5.10	3.80	3.64	4.27	3.64	4.78	3.99	3.93	4.02	3.77	7.29	6.17	6.17	5.98	6.08
		3	5.47	4.09	3.92	4.55	3.92	5.19	4.35	4.28	4.33	4.11	7.50	6.35	6.34	6.15	6.25
		4	2.23	1.72	1.62	1.77	1.64	2.29	1.97	1.93	1.79	1.86	1.79	1.50	1.50	1.47	1.48
		5	5.16	3.86	3.69	4.32	3.69	4.86	4.07	4.00	4.08	3.85	7.30	6.18	6.17	5.98	6.09
	10	1	2.10	1.81	1.79	1.89	1.77	2.29	2.08	2.07	2.06	2.03	2.06	1.83	1.83	1.85	1.81
		2	5.51	4.59	4.54	5.00	4.50	5.52	4.89	4.87	5.05	4.77	8.55	7.67	7.66	7.66	7.60
		3	5.88	4.92	4.86	5.34	4.82	5.96	5.30	5.29	5.44	5.17	8.79	7.88	7.88	7.87	7.81
		4	2.28	1.97	1.94	2.04	1.93	2.51	2.30	2.30	2.25	2.25	2.10	1.87	1.87	1.88	1.86
		5	5.56	4.65	4.59	5.05	4.56	5.60	4.97	4.96	5.12	4.85	8.56	7.68	7.67	7.67	7.61
(22) (b)	1	1	4.56	4.36	4.25	1.10	3.27	7.75	7.71	7.66	1.56	4.47	4.64	4.62	4.63	1.46	3.63
		2	10.14	9.95	9.82	2.68	7.83	13.74	13.75	13.70	3.00	8.53	10.43	10.38	10.39	3.10	8.13
		3	11.10	10.84	10.68	2.90	8.47	15.72	15.71	15.64	3.35	9.61	11.42	11.36	11.37	3.40	8.90
		4	5.18	4.92	4.78	1.21	3.65	9.53	9.47	9.39	1.76	5.31	5.30	5.24	5.24	1.53	4.11
		5	10.14	9.94	9.80	2.69	7.80	14.12	14.12	14.05	3.05	8.59	10.87	10.76	10.77	3.14	8.42
	2	1	4.89	4.64	4.55	1.38	3.56	8.15	8.09	8.06	1.81	4.86	4.46	4.47	4.47	1.63	3.51
		2	10.82	10.57	10.45	3.27	8.44	14.42	14.40	14.38	3.46	9.16	9.98	10.00	10.01	3.45	7.83
		3	11.86	11.52	11.39	3.55	9.15	16.52	16.47	16.44	3.87	10.34	10.93	10.95	10.96	3.79	8.57
		4	5.55	5.24	5.13	1.53	3.98	10.04	9.94	9.89	2.06	5.79	5.02	5.00	5.01	1.70	3.91
		5	10.83	10.55	10.44	3.28	8.41	14.84	14.79	14.76	3.53	9.26	10.32	10.30	10.31	3.49	8.05
	3	1	5.02	4.81	4.72	1.55	3.75	8.34	8.30	8.26	1.91	5.10	4.28	4.31	4.33	1.59	3.37
		2	11.12	10.90	10.79	3.49	8.80	14.75	14.76	14.73	3.54	9.52	9.57	9.65	9.69	3.38	7.52
		3	12.18	11.90	11.76	3.82	9.56	16.90	16.88	16.85	3.99	10.77	10.49	10.57	10.61	3.71	8.23
		4	5.70	5.43	5.31	1.75	4.20	10.27	10.19	10.14	2.23	6.12	4.78	4.79	4.81	1.67	3.73
		5	11.12	10.90	10.77	3.51	8.78	15.17	15.15	15.12	3.64	9.65	9.84	9.88	9.91	3.42	7.68
	6	1	5.26	5.07	4.99	1.89	4.09	8.67	8.63	8.59	2.38	5.58	3.84	3.90	3.95	1.82	3.05
		2	11.73	11.55	11.44	4.13	9.51	15.41	15.43	15.39	4.23	10.30	8.57	8.73	8.82	3.92	6.79
		3	12.85	12.60	12.47	4.54	10.34	17.65	17.65	17.60	4.83	11.70	9.38	9.55	9.66	4.29	7.43
		4	5.95	5.71	5.60	2.14	4.58	10.61	10.55	10.48	2.85	6.72	4.16	4.23	4.28	1.95	3.28
		5	11.72	11.53	11.41	4.16	9.48	15.80	15.80	15.75	4.38	10.47	8.66	8.80	8.90	3.99	6.83
	10	1	5.25	5.22	5.21	2.20	4.38	8.90	8.87	8.89	3.01	6.11	3.58	3.59	3.60	2.21	2.83
		2	12.04	12.11	12.12	4.83	10.24	16.06	16.08	16.12	5.25	11.24	7.98	8.02	8.05	4.82	6.29
		3	13.11	13.16	13.17	5.29	11.12	18.32	18.32	18.36	6.02	12.75	8.74	8.78	8.82	5.28	6.89
		4	5.87	5.82	5.81	2.47	4.88	10.77	10.71	10.73	3.62	7.33	3.80	3.79	3.80	2.43	2.96
		5	12.00	12.06	12.07	4.84	10.20	16.36	16.36	16.40	5.43	11.40	7.98	7.99	8.02	4.93	6.26

Table 6: Precision of Final Data: Weighted Root Mean Square Error (i.e., $WRMSE_{ij}^{(L)}$) — Continued

ME	L	TE	(28)					(29)					(30)				
		$i\backslash j$	(i) 1	(ii) 2	(iii) 3	(iv) 4	(v) 5	(i) 1	(ii) 2	(iii) 3	(iv) 4	(v) 5	(i) 1	(ii) 2	(iii) 3	(iv) 4	(v) 5
		1	1.90	1.89	2.08	0.83	0.89	2.61	2.50	2.57	0.97	0.99	3.22	3.06	3.07	1.48	1.36
		2	2.90	2.79	3.03	1.40	1.36	3.51	3.30	3.39	1.40	1.35	3.97	3.78	3.80	1.73	1.62
	1	3	3.42	3.33	3.62	1.60	1.61	4.27	4.05	4.16	1.66	1.65	4.99	4.78	4.80	2.11	2.00
		4	2.52	2.58	2.87	1.01	1.18	3.92	3.82	3.96	1.31	1.45	5.45	5.22	5.25	2.09	2.04
		5	3.42	3.38	3.71	1.53	1.59	4.65	4.47	4.62	1.67	1.75	5.96	5.72	5.75	2.29	2.24
		1	1.75	1.68	1.84	1.02	0.82	2.33	2.19	2.25	1.12	0.93	2.99	2.83	2.84	1.79	1.56
		2	2.73	2.51	2.72	1.68	1.27	3.18	2.94	3.00	1.60	1.29	3.65	3.47	3.48	2.07	1.81
	2	3	3.22	2.99	3.25	1.92	1.50	3.88	3.61	3.69	1.90	1.57	4.59	4.38	4.39	2.49	2.19
		4	2.27	2.26	2.51	1.26	1.06	3.40	3.29	3.40	1.49	1.33	4.89	4.69	4.71	2.41	2.13
		5	3.16	3.00	3.29	1.86	1.47	4.11	3.90	4.01	1.90	1.63	5.35	5.13	5.16	2.63	2.33
		1	1.64	1.55	1.70	1.11	0.83	2.09	1.97	2.04	1.16	0.91	2.70	2.56	2.60	1.72	1.45
(22)		2	2.59	2.35	2.53	1.77	1.28	2.90	2.67	2.74	1.64	1.26	3.30	3.15	3.20	2.02	1.70
	3	3	3.04	2.79	3.02	2.06	1.52	3.54	3.29	3.38	1.96	1.54	4.17	3.99	4.06	2.48	2.09
(c)		4	2.12	2.07	2.29	1.42	1.08	3.00	2.91	3.04	1.61	1.30	4.42	4.26	4.35	2.47	2.09
		5	2.97	2.78	3.03	2.02	1.49	3.68	3.49	3.62	2.01	1.60	4.84	4.67	4.76	2.70	2.29
		1	1.48	1.28	1.37	1.33	0.89	1.85	1.63	1.72	1.42	1.04	2.43	2.25	2.38	1.96	1.60
		2	2.41	2.02	2.13	2.09	1.39	2.63	2.28	2.38	1.96	1.44	2.96	2.75	2.92	2.32	1.90
	6	3	2.80	2.38	2.51	2.44	1.64	3.18	2.78	2.91	2.37	1.76	3.72	3.48	3.69	2.87	2.35
		4	1.86	1.65	1.79	1.74	1.17	2.54	2.31	2.48	2.02	1.51	3.91	3.69	3.93	2.90	2.38
		5	2.70	2.32	2.47	2.41	1.61	3.22	2.86	3.03	2.47	1.83	4.28	4.03	4.30	3.18	2.60
		1	1.43	1.14	1.10	1.54	1.04	1.82	1.49	1.41	1.79	1.36	2.41	2.15	2.17	2.36	1.98
		2	2.36	1.86	1.80	2.41	1.62	2.63	2.14	2.02	2.46	1.88	2.93	2.63	2.65	2.85	2.40
	10	3	2.74	2.18	2.11	2.82	1.91	3.17	2.59	2.44	3.00	2.30	3.69	3.32	3.36	3.56	3.01
		4	1.76	1.44	1.39	2.00	1.35	2.47	2.05	1.94	2.56	1.97	3.86	3.50	3.55	3.73	3.16
		5	2.62	2.10	2.03	2.79	1.88	3.16	2.61	2.47	3.13	2.41	4.22	3.83	3.88	4.08	3.46
		1	1.31	0.80	0.67	0.93	0.74	1.28	0.88	0.78	1.02	0.81	1.16	0.76	0.74	1.02	0.74
		2	3.43	1.95	1.70	2.81	1.81	3.20	2.03	1.85	2.86	1.88	4.82	3.20	3.12	4.32	3.13
	1	3	3.66	2.10	1.82	2.95	1.94	3.43	2.20	2.00	3.03	2.04	4.95	3.29	3.21	4.43	3.22
		4	1.43	0.89	0.74	1.00	0.82	1.41	0.99	0.87	1.10	0.91	1.17	0.77	0.75	1.04	0.76
		5	3.47	1.98	1.73	2.83	1.84	3.25	2.08	1.89	2.89	1.93	4.82	3.20	3.13	4.32	3.14
		1	1.63	1.05	0.96	1.18	0.96	1.58	1.16	1.10	1.20	1.05	1.43	0.98	0.97	1.19	0.96
		2	4.21	2.57	2.39	3.43	2.38	3.90	2.67	2.57	3.32	2.46	5.94	4.12	4.09	5.00	4.03
	2	3	4.50	2.76	2.57	3.62	2.56	4.20	2.90	2.79	3.53	2.67	6.10	4.23	4.20	5.14	4.14
		4	1.79	1.16	1.06	1.28	1.07	1.76	1.30	1.24	1.30	1.18	1.45	1.00	0.99	1.21	0.98
		5	4.27	2.61	2.43	3.46	2.42	3.97	2.73	2.63	3.36	2.51	5.94	4.12	4.10	5.00	4.03
		1	1.75	1.25	1.16	1.33	1.17	1.69	1.38	1.33	1.30	1.28	1.49	1.16	1.15	1.22	1.14
(23)		2	4.41	3.01	2.83	3.65	2.85	4.05	3.13	3.04	3.43	2.94	6.19	4.89	4.86	5.15	4.81
	3	3	4.73	3.25	3.05	3.88	3.07	4.38	3.41	3.32	3.66	3.20	6.36	5.02	4.99	5.29	4.94
(a)		4	1.92	1.39	1.28	1.46	1.30	1.89	1.55	1.50	1.43	1.44	1.51	1.19	1.18	1.25	1.17
		5	4.47	3.07	2.88	3.69	2.90	4.12	3.21	3.12	3.47	3.01	6.20	4.90	4.87	5.16	4.82
		1	2.04	1.57	1.49	1.63	1.49	2.05	1.76	1.72	1.62	1.66	1.76	1.47	1.47	1.44	1.45
		2	5.10	3.79	3.62	4.29	3.62	4.80	3.99	3.92	4.07	3.77	7.32	6.18	6.17	6.05	6.09
	6	3	5.48	4.09	3.91	4.58	3.91	5.21	4.35	4.28	4.38	4.11	7.53	6.35	6.34	6.22	6.26
		4	2.25	1.74	1.64	1.78	1.65	2.29	1.98	1.93	1.79	1.86	1.79	1.50	1.50	1.47	1.48
		5	5.17	3.85	3.68	4.35	3.68	4.90	4.08	4.01	4.14	3.86	7.33	6.18	6.18	6.06	6.09
		1	2.11	1.82	1.80	1.90	1.79	2.29	2.08	2.07	2.06	2.03	2.06	1.83	1.83	1.85	1.81
		2	5.52	4.58	4.53	5.02	4.49	5.55	4.90	4.88	5.10	4.78	8.59	7.70	7.69	7.72	7.63
	10	3	5.90	4.92	4.86	5.36	4.82	5.99	5.31	5.30	5.49	5.18	8.83	7.91	7.90	7.93	7.84
		4	2.30	1.99	1.96	2.06	1.95	2.51	2.30	2.30	2.26	2.25	2.10	1.87	1.87	1.88	1.86
		5	5.58	4.65	4.59	5.08	4.55	5.64	4.99	4.97	5.17	4.87	8.60	7.71	7.70	7.73	7.64

Table 6: Precision of Final Data: Weighted Root Mean Square Error (i.e., $WRMSE_{ij}^{(L)}$) — Continued

ME	L	TE	(28)					(29)					(30)				
			(i)	(ii)	(iii)	(iv)	(v)	(i)	(ii)	(iii)	(iv)	(v)	(i)	(ii)	(iii)	(iv)	(v)
		$i \backslash j$	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
(23) (b)	1	1	6.03	5.81	5.68	1.15	4.15	10.21	10.16	10.10	1.64	5.54	7.24	7.21	7.20	1.57	5.84
		2	12.77	12.54	12.37	2.79	9.50	17.44	17.43	17.37	3.16	10.28	17.31	17.28	17.26	3.68	14.00
		3	14.06	13.77	13.57	3.01	10.33	20.08	20.05	19.97	3.52	11.62	18.74	18.71	18.69	3.97	15.15
		4	6.96	6.68	6.52	1.27	4.71	12.87	12.80	12.71	1.88	6.71	7.42	7.40	7.39	1.59	5.99
		5	12.93	12.68	12.50	2.81	9.56	18.40	18.38	18.30	3.24	10.58	17.31	17.29	17.27	3.68	14.01
	2	1	6.34	6.08	5.97	1.45	4.46	10.60	10.52	10.48	1.94	5.96	7.62	7.55	7.55	1.83	6.18
		2	13.44	13.15	13.01	3.43	10.14	18.12	18.08	18.04	3.69	10.96	18.19	18.08	18.08	4.26	14.81
		3	14.80	14.44	14.28	3.71	11.05	20.89	20.82	20.77	4.12	12.43	19.70	19.58	19.59	4.61	16.04
		4	7.31	6.98	6.85	1.63	5.06	13.35	13.23	13.17	2.24	7.24	7.80	7.74	7.74	1.85	6.34
		5	13.60	13.29	13.14	3.45	10.21	19.10	19.04	18.99	3.79	11.31	18.19	18.08	18.09	4.26	14.81
	3	1	6.45	6.22	6.11	1.64	4.66	10.76	10.70	10.65	2.06	6.22	7.77	7.73	7.72	1.83	6.37
		2	13.71	13.47	13.33	3.66	10.52	18.42	18.40	18.36	3.80	11.37	18.58	18.52	18.52	4.29	15.27
		3	15.10	14.79	14.62	4.00	11.48	21.23	21.19	21.14	4.28	12.91	20.12	20.06	20.05	4.64	16.54
		4	7.42	7.13	6.99	1.86	5.29	13.52	13.42	13.36	2.44	7.59	7.97	7.93	7.93	1.86	6.54
		5	13.87	13.61	13.46	3.70	10.60	19.40	19.36	19.31	3.93	11.75	18.59	18.53	18.52	4.30	15.28
	6	1	6.58	6.38	6.28	2.00	4.98	10.92	10.87	10.81	2.56	6.73	8.15	8.07	8.06	2.18	6.73
		2	14.20	14.01	13.88	4.32	11.24	18.93	18.93	18.88	4.53	12.20	19.48	19.35	19.33	5.12	16.13
		3	15.59	15.35	15.19	4.74	12.26	21.76	21.74	21.67	5.16	13.88	21.10	20.95	20.94	5.55	17.46
		4	7.52	7.26	7.14	2.28	5.64	13.59	13.51	13.42	3.11	8.21	8.35	8.28	8.27	2.22	6.90
		5	14.33	14.13	13.99	4.38	11.32	19.81	19.79	19.72	4.73	12.62	19.48	19.35	19.34	5.13	16.13
	10	1	6.40	6.36	6.36	2.33	5.22	10.91	10.87	10.89	3.22	7.24	8.47	8.39	8.39	2.77	7.09
		2	14.32	14.39	14.40	5.06	11.95	19.34	19.35	19.38	5.60	13.16	20.27	20.13	20.14	6.50	17.01
		3	15.66	15.71	15.72	5.56	13.02	22.13	22.12	22.16	6.41	14.97	21.94	21.79	21.80	7.04	18.42
		4	7.21	7.16	7.15	2.63	5.86	13.37	13.30	13.32	3.93	8.79	8.68	8.61	8.61	2.83	7.28
		5	14.41	14.47	14.48	5.11	12.01	20.05	20.04	20.07	5.84	13.57	20.28	20.14	20.14	6.50	17.02
(23) (c)	1	1	1.79	1.78	1.99	0.80	0.80	2.49	2.38	2.47	0.94	0.92	3.10	2.94	2.95	1.42	1.30
		2	2.72	2.59	2.84	1.34	1.23	3.33	3.11	3.20	1.34	1.25	3.78	3.59	3.61	1.66	1.54
		3	3.21	3.10	3.41	1.53	1.45	4.05	3.82	3.94	1.60	1.53	4.76	4.54	4.56	2.03	1.90
		4	2.41	2.48	2.79	0.98	1.08	3.83	3.73	3.90	1.28	1.38	5.37	5.14	5.17	2.06	2.00
		5	3.20	3.16	3.50	1.46	1.44	4.46	4.28	4.44	1.61	1.63	5.80	5.55	5.58	2.24	2.16
	2	1	1.65	1.57	1.75	0.97	0.75	2.22	2.08	2.15	1.08	0.88	2.87	2.70	2.71	1.72	1.49
		2	2.56	2.32	2.54	1.60	1.16	3.01	2.75	2.82	1.53	1.21	3.47	3.27	3.29	1.98	1.73
		3	3.02	2.78	3.05	1.84	1.37	3.67	3.39	3.48	1.82	1.47	4.37	4.14	4.16	2.39	2.10
		4	2.16	2.16	2.43	1.21	0.99	3.30	3.19	3.33	1.46	1.28	4.79	4.58	4.60	2.36	2.10
		5	2.95	2.79	3.10	1.77	1.35	3.92	3.70	3.84	1.83	1.54	5.18	4.94	4.97	2.56	2.27
	3	1	1.54	1.45	1.61	1.06	0.77	1.98	1.85	1.94	1.12	0.87	2.57	2.43	2.47	1.65	1.39
		2	2.42	2.16	2.36	1.68	1.19	2.74	2.49	2.57	1.56	1.20	3.11	2.95	3.00	1.93	1.64
		3	2.85	2.59	2.83	1.95	1.41	3.34	3.07	3.17	1.87	1.46	3.95	3.76	3.82	2.37	2.02
		4	2.00	1.96	2.21	1.37	1.02	2.88	2.79	2.95	1.57	1.26	4.29	4.13	4.21	2.42	2.08
		5	2.77	2.57	2.84	1.91	1.38	3.48	3.28	3.43	1.92	1.52	4.65	4.46	4.55	2.62	2.25
	6	1	1.39	1.19	1.29	1.27	0.86	1.76	1.53	1.63	1.36	1.01	2.32	2.13	2.25	1.87	1.56
		2	2.26	1.86	1.97	1.97	1.31	2.50	2.13	2.22	1.86	1.38	2.80	2.58	2.73	2.20	1.84
		3	2.63	2.19	2.33	2.31	1.55	3.02	2.59	2.72	2.25	1.69	3.53	3.27	3.47	2.73	2.28
		4	1.75	1.54	1.70	1.68	1.14	2.43	2.18	2.38	1.96	1.50	3.78	3.53	3.78	2.84	2.37
		5	2.52	2.12	2.29	2.28	1.53	3.04	2.67	2.85	2.36	1.78	4.10	3.83	4.09	3.07	2.57
	10	1	1.36	1.07	1.03	1.48	1.01	1.75	1.41	1.33	1.72	1.34	2.32	2.05	2.06	2.26	1.94
		2	2.24	1.74	1.67	2.28	1.56	2.53	2.02	1.89	2.33	1.83	2.81	2.49	2.50	2.70	2.33
		3	2.60	2.03	1.96	2.68	1.85	3.04	2.45	2.29	2.85	2.24	3.55	3.16	3.17	3.39	2.93
		4	1.68	1.36	1.30	1.94	1.34	2.39	1.97	1.85	2.50	1.98	3.78	3.39	3.41	3.65	3.17
		5	2.47	1.95	1.87	2.63	1.82	3.03	2.47	2.31	2.99	2.36	4.10	3.68	3.70	3.95	3.42

small. When we take an example of the case: (22).(a), (28), $L = 1$ and (i), we have 1.29 for $i = 1$, 3.40 for $i = 2$, 3.62 for $i = 3$, 1.41 for $i = 4$ and 3.43 for $i = 5$. Therefore, in this case, the predictor $i = 3$, i.e., (iii), has the smallest variance, because $WRMSE_{i1}^{(1)}$ is the largest of $i = 1, \cdots, 5$. For all the cases in Table 6, we obtain the result that the predictor (iii) has the smallest variance.

5 SUMMARY

In this paper, we have dealt with estimation of final data given revised data and/or preliminary data. The Monte-Carlo integration nonlinear filter proposed by Tanizaki (1991,1993) and Tanizaki and Mariano (1994) was applied. The predictors (i) – (v) were introduced as an application of the nonlinear filtering problem, i.e., (i) $E(x_t^f|X_{t-1}^p)$, (ii) $E(x_t^f|X_t^p)$, (iii) $E(x_t^f|X_T^p)$, (iv) $E(x_t^f|_tX_{t-1}^f, X_{t-1}^p)$, and (v) $E(x_t^f|_tX_{t-1}^f, X_t^p)$. In Sections 3.1 and 3.2, the algorithms for the optimal predictors in the sense of minimum mean square error were derived in the following two cases: linear and normal case, and nonlinear and/or nonnormal case. The former gives the linear algorithms while the latter utilizes the weight functions with the Monte-Carlo integration.

The U.S. consumption function was taken as an empirical example, where the measurement equation was simply assumed to be linear or log-linear between the preliminary data and the final data while the transition equation was derived from the utility maximization problem. We have considered six measurement equations (i.e., (a) – (c) for each of (22) and (23)) and three transition equations (i.e., (28) – (30)). The following results have been obtained.

1. From $MAPE$, (22).(a) or (23).(a) for the measurement equation and (29) for the transition equation gives us the best assumption.
2. (v) is the best predictor from the criteria $MAPE$ and $WRMSE$ but not too different from (iii). Recall that (v) utilizes more information set than the predictors (i), (ii) and (iv).
3. For (i) – (iii) of model (c), $WRMSE$ decreases as L increases. However, for the other, $WRMSE$ increases as L increases. Accordingly, except for (i) – (iii) of (c), discrepancy between the L-th revised data and the estimated final data becomes large as L is large.
4. The predictor (iii) gives us the smallest variance from $WRMSE$.

In general, the final data is unobservable because it is revised forever. It might be plausible to consider that the discrepancy between the L-th revised data and the final data disappears as L increases (i.e., as the data is revised). In this sense, we compared the predictors (i) – (v) with the L-th revised data for large L. As a result, (iii) of model (c) is recommended.

Also, sometimes, it is useful for economic forecasting to estimate the first (or second) revised data. Therefore, since the difference between the predictor (v) of model (a) and the L-th revised data for small L is the smallest, (v) of model (a) is recommended in this case.

APPENDIX: Monte-Carlo Integration

Let x be a random variable with distribution $P(x)$ and $z(x)$ be a function of x. The expectation of $z(x)$ is defined as follows.

$$E\big(z(x)\big) = \int z(x)P(x)\mathrm{d}x.$$

When the above integration cannot be evaluated explicitly, Monte-Carlo integration is often used. The above expectation can be rewritten as $E\big(z(x)\big) = \int z(x)\omega(x)P_x(x)\mathrm{d}x$, where $\omega(x) = P(x)/P_x(x)$. Generating n random draws from the importance density $P_x(x)$, the expectation is approximated. Let $\overline{z}_n$ be the estimate of $E\big(z(x)\big)$, which is computed as follows.[23]

$$\overline{z}_n = \frac{1}{n}\sum_{i=1}^{n} z(x_i)\omega(x_i), \tag{31}$$

where x_i is a random number generated from $P_x(x)$.

Now, we take an example of $z(x) = (x-\mu)/\sigma$, where the underlying distribution of x is normal with mean μ and variance σ^2, i.e., $P(x) = N(\mu, \sigma^2)$. Note that $z(x)$ has a standard normal distribution, i.e., $z(x) \sim N(0,1)$.

We have chosen the bimodal distribution function for the importance density in the empirical example of Section 4, which is the importance density $P_x(x_t^f)$ such that $P_x(x_t^f)$ covers both $P(x_{t-1}^f|X_{t-1}^p)$ and $P(x_{t-1}^f|X_t^p)$, Therefore, in this appendix, we examine the importance density which is assumed to be: $P_x(x) = \frac{1}{2}N(\mu_1,\sigma_1^2) + \frac{1}{2}N(\mu_2,\sigma_2^2)$, where $\mu_1 = -4$, $\mu_2 = 4$, $\sigma_1 = 1$ and $\sigma_2 = 4$. The number of random draws is $n = 50,\ 100,\ 200$. Moreover, we choose the parameters μ and σ as $\mu = \mu_i + c_\mu\sigma_i$ for $c_\mu = \pm 2.0,\ \pm 1.5,\ \pm 1.0,\ \pm 0.5,\ 0.0$ and $i = 1,2$ and $\sigma = c_\sigma\sigma_i$ for $c_\sigma = 1/3,\ 1/2,\ 2/3,\ 1,\ 3/2,\ 2,\ 3$ and $i = 1,2$.

Moreover, in order to examine precision of Monte-Carlo integration, we repeat the simulation run K times and obtain K estimates of $\overline{z}_n$, where $K = 10,000$ is taken. Thus, K simulation runs are performed. For each μ and σ, we compute the sample mean of $\overline{z}_n$ in Table 7 and the sample standard error of $\overline{z}_n$ in Table 8.

The asymptotic properties are as follows.[24] Note that $z(x_i)\omega(x_i)$, $i = 1,\cdots,n$, are random. Define $m = E\big(z(x_i)\omega(x_i)\big) = E\big(z(x)\big)$ and $s^2 = Var\big(z(x_i)\omega(x_i)\big) = E\big(z(x_i)\omega(x_i)\big)^2 - m^2 = E\Big(\big(z(x)\big)^2\omega(x)\Big) - m^2$. By the central limit theorem, we can easily show that

$$\frac{\overline{z}_n - m}{s/\sqrt{n}} \longrightarrow N(0,1),$$

where $m = 0$ in this example. Thus, $\overline{z}_n$ is consistent. For any μ and σ^2, $\overline{z}_n$ gives us an asymptotically unbiased estimator as n goes to infinity. However, convergence is quite slow as $\sqrt{n}$.

In Table 7, the sample means of $\overline{z}_n$ are close to zero except for large σ or large μ in absolute value, which implies that $\overline{z}_n$ is asymptotically unbiased. In Table 8, the sample variances of $\overline{z}_n$ decrease as n is large. Moreover, around $\mu = -4,\ 4$ and $\sigma = 0.5 \sim 1,\ 2 \sim 4$, the sample variances of $\overline{z}_n$ are small.

When either of two peaks of $P_x(x)$ is close to center of $P(x)$, $\overline{z}_n$ is asymptotically unbiased. And when $P_x(x)$ are wider in range than $P(x)$, the sample variance of $\overline{z}_n$ is small. Therefore, it might be concluded from the simulation study that we should choose the importance density with the following conditions: (i) the importance density should have a wide range of distribution, compared with the original distribution and (ii) center of the importance density should be close to that of the original density but we do not have to pay too much attention to center of the importance density.

The importance density used in Section 3:2.(i) and (ii) (i.e., $P_x(x_t^f)$) needs to cover $P(x_t^f|X_s^p)$, $s = t-1, t$. Peak and range of $P(x_t^f|X_{t-1}^p)$ are different from those of $P(x_t^f|X_t^p)$. In general, range of $P(x_t^f|X_{t-1}^p)$ is larger than that of $P(x_t^f|X_t^p)$. For the Monte-Carlo integration filter, two densities $P(x_t^f|X_s^p)$, $s = t-1, t$, have to be approximated by one importance density $P_x(x_t^f)$. Thus, the simulation study in this appendix shows that the bimodal importance density is appropriate for the Monte-Carlo integration filter.

[23] Precision of the following approximation is much better than equation (31).

$$\overline{z}_n = \sum_{i=1}^{n} z(x_i)\omega(x_i) \Big/ \sum_{i=1}^{n}\omega(x_i),$$

where $\sum_{i=1}^{n} z(x_i)\omega(x_i)$ is divided by $\sum_{i=1}^{n}\omega(x_i)$, not n. In this paper, all the results related to Monte-Carlo integration (i.e., Tables 4 – 6, 7 and 8) are computed using the above approximation.

[24] Note that x_i is a random variable from $P_x(x)$ while x is from $P(x)$. Therefore, $E\big(z(x_i)\omega(x_i)\big)$ implies taking the expectation with respect to x_i but $E\big(z(x)\big)$ is with respect to x.

Table 7: Precision of Monte-Carlo Integration — Sample Mean of $\bar{z}_n$

μ	n \ σ	0.33	0.5	0.67	1	1.33	1.5	2	2.67	3	4	6	8	12
	50	.534	.410	.382	.397	.428	.444	.489	.534	.552	.590	.638	.662	.611
−6.0	100	.186	.173	.186	.232	.278	.300	.363	.431	.458	.517	.586	.623	.598
	200	.070	.076	.091	.129	.171	.192	.253	.332	.366	.443	.533	.582	.582
	50	.119	.112	.129	.184	.245	.273	.347	.422	.451	.514	.587	.622	.579
−5.5	100	.041	.046	.058	.096	.145	.170	.243	.329	.364	.444	.536	.584	.567
	200	.018	.022	.029	.051	.084	.102	.161	.244	.283	.375	.484	.545	.552
	50	.035	.035	.041	.071	.121	.149	.229	.321	.358	.442	.537	.582	.546
−5.0	100	.016	.017	.020	.035	.065	.084	.149	.240	.281	.376	.487	.546	.536
	200	.008	.008	.010	.019	.036	.049	.095	.170	.210	.312	.438	.508	.522
	50	.012	.012	.014	.025	.050	.068	.137	.233	.275	.374	.490	.544	.514
−4.5	100	.005	.006	.007	.012	.024	.035	.082	.165	.207	.313	.441	.509	.505
	200	.003	.004	.004	.007	.014	.020	.051	.113	.150	.255	.394	.473	.492
	50	−.001	.001	.002	.005	.014	.023	.070	.158	.203	.313	.444	.506	.482
−4.0	100	.000	.000	.001	.002	.006	.010	.039	.106	.146	.256	.397	.473	.474
	200	.001	.001	.001	.001	.004	.006	.024	.070	.102	.204	.353	.438	.463
	50	−.007	−.007	−.007	−.008	−.007	−.003	.025	.099	.142	.256	.400	.469	.450
−3.5	100	−.003	−.004	−.005	−.005	−.005	−.003	.012	.062	.097	.204	.356	.438	.444
	200	−.001	−.002	−.003	−.003	−.002	−.001	.009	.041	.065	.159	.313	.405	.433
	50	−.021	−.019	−.020	−.023	−.022	−.020	−.002	.055	.093	.206	.359	.433	.418
−3.0	100	−.012	−.011	−.012	−.013	−.012	−.011	−.002	.032	.059	.159	.316	.403	.413
	200	−.007	−.008	−.007	−.007	−.006	−.005	.000	.021	.040	.121	.276	.372	.404
	50	−.058	−.050	−.046	−.041	−.035	−.031	−.016	.024	.055	.161	.320	.397	.386
−2.5	100	−.023	−.023	−.022	−.020	−.017	−.016	−.009	.013	.033	.121	.279	.370	.382
	200	−.013	−.012	−.011	−.010	−.008	−.007	−.004	.009	.022	.090	.242	.341	.374
	50	−.156	−.099	−.077	−.054	−.041	−.035	−.022	.006	.029	.123	.283	.362	.354
−2.0	100	−.049	−.037	−.032	−.024	−.019	−.017	−.012	.002	.016	.089	.245	.337	.352
	200	−.015	−.013	−.012	−.011	−.009	−.008	−.005	.003	.011	.064	.210	.310	.345
	50	.097	.056	.038	.019	.009	.006	.000	−.002	.000	.030	.157	.227	.227
0.0	100	.030	.019	.014	.008	.003	.002	−.001	−.002	−.001	.017	.131	.214	.230
	200	.010	.007	.005	.003	.001	.000	−.001	−.001	.000	.013	.109	.197	.229
	50	.025	.016	.013	.013	.013	.013	.011	.007	.005	.004	.061	.099	.100
2.0	100	.007	.005	.005	.005	.006	.006	.005	.004	.003	.002	.052	.099	.109
	200	.006	.003	.003	.002	.003	.003	.003	.002	.002	.003	.044	.095	.113
	50	.002	.000	.001	.002	.002	.002	.002	.001	.000	−.005	−.020	−.028	−.028
4.0	100	.001	.001	.002	.002	.003	.003	.002	.002	.001	−.002	−.007	−.011	−.012
	200	.001	.001	.002	.002	.002	.003	.003	.002	.002	.000	.000	−.001	−.002
	50	−.013	−.010	−.011	−.011	−.011	−.010	−.011	−.013	−.015	−.029	−.109	−.157	−.155
6.0	100	−.010	−.006	−.005	−.005	−.004	−.004	−.004	−.006	−.007	−.013	−.071	−.124	−.134
	200	−.003	−.001	−.001	−.001	−.001	−.001	−.001	−.002	−.002	−.005	−.046	−.098	−.117
	50	−.116	−.069	−.052	−.041	−.037	−.037	−.038	−.047	−.055	−.097	−.227	−.294	−.284
8.0	100	−.020	−.014	−.013	−.015	−.016	−.016	−.018	−.022	−.026	−.048	−.163	−.243	−.256
	200	−.007	−.008	−.008	−.008	−.008	−.008	−.009	−.011	−.012	−.024	−.117	−.203	−.233
	50	−.592	−.362	−.258	−.171	−.141	−.134	−.134	−.160	−.180	−.258	−.382	−.440	−.413
10.0	100	−.133	−.079	−.062	−.053	−.052	−.053	−.058	−.074	−.087	−.154	−.294	−.375	−.379
	200	−.037	−.026	−.023	−.023	−.025	−.025	−.028	−.036	−.042	−.089	−.227	−.321	−.350
	50	−2.593	−1.692	−1.243	−.813	−.627	−.575	−.500	−.488	−.495	−.526	−.575	−.598	−.544
12.0	100	−1.048	−.661	−.474	−.305	−.246	−.233	−.228	−.262	−.288	−.365	−.466	−.520	−.504
	200	−.312	−.188	−.136	−.095	−.087	−.088	−.099	−.133	−.159	−.248	−.379	−.454	−.469

Table 8: Precision of Monte-Carlo Integration — Sample Standard Error of $\overline{z}_n$

μ	n \ σ	0.33	0.5	0.67	1	1.33	1.5	2	2.67	3	4	6	8	12
	50	1.066	.813	.714	.625	.579	.561	.506	.436	.408	.348	.280	.256	.240
−6.0	100	.653	.551	.531	.533	.535	.535	.518	.466	.438	.374	.300	.270	.248
	200	.383	.354	.368	.415	.452	.465	.492	.476	.454	.392	.313	.276	.249
	50	.570	.476	.465	.492	.504	.504	.488	.436	.410	.350	.282	.260	.242
−5.5	100	.326	.298	.312	.378	.431	.450	.477	.456	.434	.375	.302	.272	.250
	200	.211	.200	.213	.274	.339	.365	.425	.451	.440	.388	.313	.277	.251
	50	.330	.284	.283	.341	.406	.427	.455	.430	.408	.352	.284	.263	.244
−5.0	100	.212	.188	.190	.243	.321	.354	.421	.438	.425	.374	.303	.275	.252
	200	.145	.130	.133	.172	.238	.271	.352	.415	.418	.383	.313	.279	.252
	50	.245	.208	.199	.230	.304	.340	.408	.418	.403	.352	.286	.267	.246
−4.5	100	.168	.144	.137	.160	.225	.263	.355	.410	.410	.371	.303	.278	.253
	200	.116	.100	.096	.113	.162	.193	.281	.369	.388	.374	.312	.280	.254
	50	.223	.188	.175	.183	.231	.264	.352	.398	.393	.351	.288	.272	.248
−4.0	100	.153	.132	.123	.129	.165	.194	.288	.372	.387	.366	.304	.281	.255
	200	.106	.092	.086	.090	.117	.139	.218	.316	.350	.363	.310	.282	.255
	50	.240	.201	.186	.186	.203	.221	.296	.368	.376	.348	.289	.276	.250
−3.5	100	.162	.139	.131	.131	.144	.158	.228	.326	.355	.358	.303	.284	.257
	200	.114	.099	.093	.092	.101	.111	.167	.263	.306	.348	.308	.283	.257
	50	.301	.250	.231	.217	.209	.209	.248	.330	.351	.343	.290	.280	.251
−3.0	100	.202	.173	.161	.151	.145	.146	.182	.276	.316	.346	.303	.287	.258
	200	.138	.121	.112	.105	.101	.102	.130	.213	.259	.329	.304	.285	.258
	50	.447	.349	.304	.255	.223	.212	.213	.287	.319	.336	.291	.285	.253
−2.5	100	.273	.226	.202	.172	.152	.145	.151	.228	.273	.332	.301	.290	.260
	200	.182	.154	.138	.118	.105	.101	.106	.169	.213	.307	.300	.286	.260
	50	.713	.492	.391	.289	.233	.214	.191	.244	.282	.325	.292	.289	.255
−2.0	100	.394	.295	.245	.188	.156	.145	.132	.185	.229	.313	.300	.293	.262
	200	.242	.192	.163	.128	.108	.100	.092	.133	.172	.281	.296	.288	.261
	50	1.176	.735	.536	.358	.274	.246	.187	.154	.161	.249	.295	.307	.261
0.0	100	.552	.378	.304	.226	.181	.164	.128	.106	.113	.208	.288	.305	.267
	200	.300	.230	.193	.150	.123	.112	.088	.074	.079	.161	.270	.294	.266
	50	.779	.502	.388	.292	.248	.234	.201	.169	.157	.171	.301	.324	.266
2.0	100	.386	.285	.238	.191	.167	.159	.139	.118	.110	.124	.275	.317	.271
	200	.239	.189	.161	.132	.117	.111	.097	.083	.077	.088	.241	.300	.270
	50	.671	.442	.349	.268	.229	.217	.192	.177	.173	.178	.318	.337	.271
4.0	100	.343	.261	.221	.179	.156	.149	.134	.125	.122	.126	.280	.326	.275
	200	.221	.177	.151	.124	.108	.103	.093	.087	.086	.088	.235	.306	.272
	50	.755	.488	.377	.285	.244	.232	.211	.205	.208	.248	.348	.346	.274
6.0	100	.382	.283	.238	.192	.167	.159	.146	.143	.145	.179	.310	.333	.277
	200	.236	.188	.161	.131	.116	.111	.102	.100	.101	.125	.266	.312	.274
	50	1.181	.750	.559	.399	.337	.319	.292	.295	.308	.375	.379	.351	.276
8.0	100	.545	.377	.309	.247	.217	.208	.194	.201	.214	.295	.349	.337	.277
	200	.302	.234	.200	.164	.147	.142	.134	.139	.148	.218	.312	.317	.274
	50	2.299	1.496	1.108	.749	.601	.561	.506	.500	.506	.489	.403	.353	.277
10.0	100	1.135	.731	.553	.405	.348	.334	.322	.347	.373	.425	.379	.339	.277
	200	.536	.376	.308	.248	.223	.217	.213	.238	.264	.353	.349	.321	.273
	50	4.106	2.740	2.067	1.416	1.112	1.018	.845	.716	.665	.547	.416	.353	.277
12.0	100	2.565	1.695	1.271	.884	.724	.679	.615	.591	.575	.503	.396	.340	.276
	200	1.352	.876	.665	.488	.427	.413	.409	.450	.467	.451	.372	.323	.271

Moreover, peak and range of $P(x_t^f|X_s^p)$ are not known, but mean and variance of x_t^f estimated by the extended Kalman filter are available. It might be plausible to consider that the true values are not too far from the extended Kalman filter estimates. Utilizing the extended Kalman filter estimates, the Monte-Carlo integration filter estimates would be improved. Therefore, in the empirical example of Section 4, we have chosen the bimodal density with $\mu_{t|1}^*$, $\mu_{t|2}^*$, $\Sigma_{t|1}^*$ and $\Sigma_{t|2}^*$, which are the extended Kalman filter estimates such that $\mu_{t|i}^* = E(x_t^f|\Omega_i)$ and $\Sigma_{t|i}^* = Var(x_t^f|\Omega_i)$ for $i = 1, 2$.

REFERENCES

Alspach, D.L. and H.W. Sorenson, 1972, "Nonlinear Bayesian Estimation Using Gaussian Sum Approximations," *IEEE Transactions on Automatic Control*, Vol. AC-17, No.4, pp.439 – 448.

Anderson, B.D.O. and J.B. Moore, 1979, *Optimal Filtering*, Prentice-Hall, New York.

Brown, B.W. and R.S. Mariano, 1984, "Residual-based Procedures for Prediction and Estimation in a Nonlinear Simultaneous System," *Econometrica*, Vol.52, No.2, pp.321 – 343.

Brown, B.W. and R.S. Mariano, 1989, "Measures of Deterministic Prediction Bias in Nonlinear Models," *International Economic Review*, Vol.30, No.3, pp.667 – 684.

Carlin, B.P., N.G. Polson and D.S. Stoffer, 1992, "A Monte Carlo Approach to Nonnormal and Nonlinear State Space Modeling," *Journal of the American Statistical Association*, Vol.87, pp.493 – 500.

Conrad, W. and C. Corrado, 1979, "Application of the Kalman Filter to Revisions in Monthly Retail Sales Estimates," *Journal of Economic Dynamic and Control*, Vol.1, pp.177 – 198.

Diebold, F.X. and M. Nerlove, 1989, "Unit Roots in Economic Time Series: A Selective Survey," in *Advances in Econometrics, Vol.8*, pp.3-69, JAI Press Inc.

Gelb, A. (Ed.), 1974, *Applied Optimal Estimation*, MIT Press.

Geweke, J., 1989, "Modeling with Normal Polynomial Expansions," in *Economic Complexity: Chaos, Sunspots, Bubbles, and Nonlinearity*, edited by Barnett et al. Cambridge University Press.

Hall, R.E., 1978, "Stochastic Implications of the Life Cycle-Permanent Income Hypothesis: Theory and Evidence," *Journal of Political Economy*, Vol.86, No.6, pp.971 – 987.

Harvey, A.C., 1989, *Forecasting, Structural Time Series Models and the Kalman Filter*, Cambridge University Press.

Howrey, E.P., 1978, "The Use of Preliminary Data in Econometric Forecasting," *The Review of Economics and Statistics*, Vol.60, pp.193 – 200.

Howrey, E.P., 1984, "Data Revision, Reconstruction, and Prediction: An Application to Inventory Investment," *The Review of Economics and Statistics*, Vol.66, pp.386 – 393.

Kramer, S.C. and Sorenson H.W., 1988, "Recursive Bayesian Estimation Using Piece-wise Constant Approximations," *Automatica*, Vol.24, No.6, pp.789 – 801.

Kitagawa, G., 1987, "Non-Gaussian State-Space Modeling of Nonstationary Time Series," *Journal of the American Statistical Association*, Vol.82, pp.1032 – 1063 (with discussion).

Kocherlakota, N.R., 1990, "On the 'Discount' Factor in Growth Economies," *Journal of Monetary Economics*, Vol.25, pp.43 – 48.

Mariano, R.S. and B.W. Brown, 1983, "Asymptotic Behavior of Predictors in a Nonlinear Simultaneous System," *International Economic Review*, Vol.24, No.3, pp.523 – 536.

Mariano, R.S. and B.W. Brown, 1989, "Stochastic Simulation, Prediction and Validation of Nonlinear Models," in *Economics in Theory and Practice: An Eclectic Approach*, edited by L.R. Klein and J. Marquez, pp.17 – 36, Kluwer Academic Publishers.

Meinhold, R.J. and N.D. Singpurwalla, 1989, "Robustification of Kalman Filter Models," *Journal of the American Statistical Association*, Vol.84, pp.479 – 486.

Shao, J., 1989, "Monte Carlo Approximations in Bayesian Decision Theory," *Journal of the American Statistical Association*, Vol.84, No.407, pp.727 – 732.

Sorenson, H.W. and D.L. Alspach, 1971, "Recursive Bayesian Estimation Using Gaussian Sums," *Automatica*, Vol.7, pp.465 – 479.

Tanizaki, H., 1991, "Nonlinear Filters: Estimation and Applications," Dissertation, University of Pennsylvania.

Tanizaki, H., 1993, *Nonlinear Filters: Estimation and Applications*, (Lecture Notes in Economics and Mathematical Systems, No.400), Springer-Verlag.

Tanizaki, H. and R.S. Mariano, 1994, "Prediction, Filtering and Smoothing in Non-linear and Non-normal Cases Using Monte Carlo Integration," *Journal of Applied Econometrics*, Vol.9, No.2, pp.163 – 179.

Wishner, R.P., J.A. Tabaczynski and M. Athans, 1969, "A Comparison of Three Non-Linear Filters," *Automatica*, Vol.5, pp.487 – 496.